Microchemical Journal

Confocal Raman spectroscopy assisted by chemometric tools: A green approach for classification and quantification of octyl p-methoxycinnamate in oil-in-water microemulsions

--Manuscript Draft--

Manuscript Number:	MICROC-D-22-02627R2
Article Type:	Research Paper
Section/Category:	Molecular spectroscopy
Keywords:	UV filters; Octyl p-methoxycinnamate; Microemulsions; Confocal Raman spectroscopy; Chemometrics; Green Analytical chemistry
Corresponding Author:	Marcos Grünhut Universidad Nacional del Sur/CONICET Bahía Blanca, Buenos Aires Argentina
First Author:	Danielle Silva do Nascimento, PhD
Order of Authors:	Danielle Silva do Nascimento, PhD
	Verónica Volpe, Pharmaceutical
	Cintia Fernandez, Chemistry
	Martín Oresti, PhD
	Lorna Ashton, PhD
	Marcos Grünhut, PhD
Abstract:	This work proposes a green analytical method based on confocal Raman spectrometry and chemometrics tools for the qualitative and quantitative analysis of oil in water microemulsions loaded with the UVB filter octyl p-methoxycinnamate (OMC). The method does not use reagents and only 10 µL of sample are needed. The analyzed microemulsion samples were synthetized in the laboratory using decaethylene glycol mono-dodecyl ether (21.9%) as non-ionic surfactant, ethyl alcohol (7.3%) as co-surfactant, oleic acid (1.5%) as oil phase and water (69.3%). A physicochemical characterization of the samples was carried out obtaining expected values for droplet size (<20 nm), polydispersity index (<0.290) and conductivity (0.04-0.07 mS cm−1), among others. Linear discriminant analysis (LDA) after selection of variables using the successive projections algorithm (SPA) and soft independent modelling of class analogy (SIMCA) were employed to classify microemulsions with different concentrations of OMC (1.0 to 10.0%). In the case of LDA, seven Raman spectral variables were previously selected by SPA and after this SPA-LDA model resulted in one error in the prediction set achieving an accuracy of 98%. The SIMCA model (α=0.05) presented an explained variance higher 97% using four principal components and it allowed the correct classification of 100% of the samples (N=15). In the quantitative analysis, partial least squares (PLS) was used to determine OMC in a range according to international legislation. The model presented optimal statistical parameters (R2 =0.9699; RMSEP=0.54%) and the prediction of samples were in close agreement with HPLC method. Moreover, the greenery of the method was estimated using the AGREE metric and an optimal value of 0.85 was obtained demonstrating the proposed analytical method results environmentally friendly.

Cover Letter

INQUISUR

November 3rd, 2022

Dear Editor,

We are sending the revised version of the manuscript titled **“Confocal Raman spectroscopy assisted by chemometric tools: A green approach for classification and quantification of octyl p-methoxycinnamate in oil-in-water microemulsions”** (Manuscript ID: MICROC-D-22-02627R1).

On this occasion, we reduced the number of figures to 4 as suggested in the author's recommendations. The changes are highlighted in green. The rest of the figures were included in the Supplementary Material.

Thank you for the very careful review of our paper, and for the comments, corrections and suggestions that ensued.

We hope that it can be accepted in your journal.

Yours sincerely,

Prof. Dr. Marcos Grünhut
Department of Chemistry-INQUISUR
Universidad Nacional del Sur-CONICET

Point by point response to reviewer's comments

Journal: Microchemical Journal

Manuscript ID: MICROC-D-22-02627 R1

Title: Confocal Raman spectroscopy assisted by chemometric tools: A green approach for classification and quantification of octyl p-methoxycinnamate in oil-in-water microemulsions

Editor and Reviewer comments:

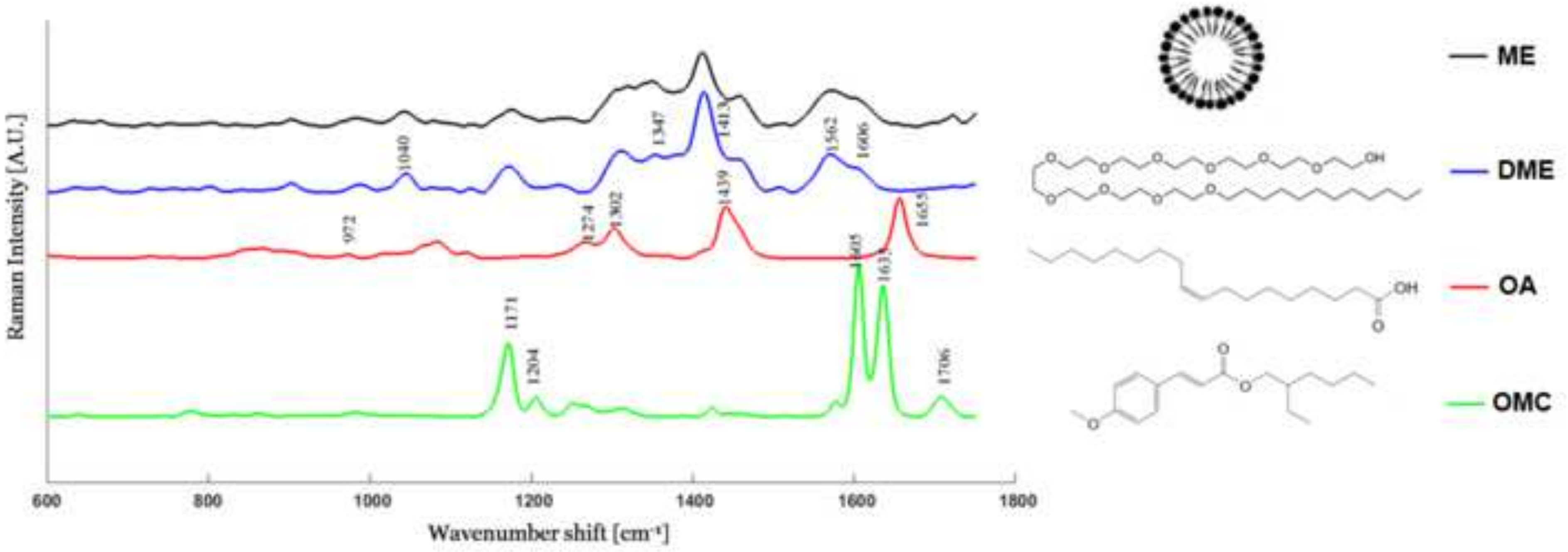
The manuscript contains an excessive number of figures not all required. So, I suggest to create a Suplementary Material section, to be included in the electronic versions but not in the printed one, and move to this section secondary figures in order to let in the main manuscript a maximum of 4 as suggested in the author's recommendations.

Response:

As requested by the Editor, the number of figures was reduced to 4. The changes are highlighted in green. The rest of the figures were included in the Supplementary Material.

Highlights

- The proposed method does not use reagents and only 10 uL of sample are used
- Only 7 spectral variables were selected by SPA to build a LDA classification model
- PLS applied to Raman spectra allowed to quantify OMC and build distribution maps
- No separation techniques were used and no waste was generated in the analysis
- A greenery value of 0.85 was obtained for the method using the AGREE metric


Raman Intensity [A.U.]
Wavenumber shift [cm⁻¹]
600
800
1000
1200
1400
1600
1800
1040
1347
1413
1562
1606
972
1274
1302
1439
1655
1171
1204
1605
1633
1706
OH
O
ME
DME
OA
OMC

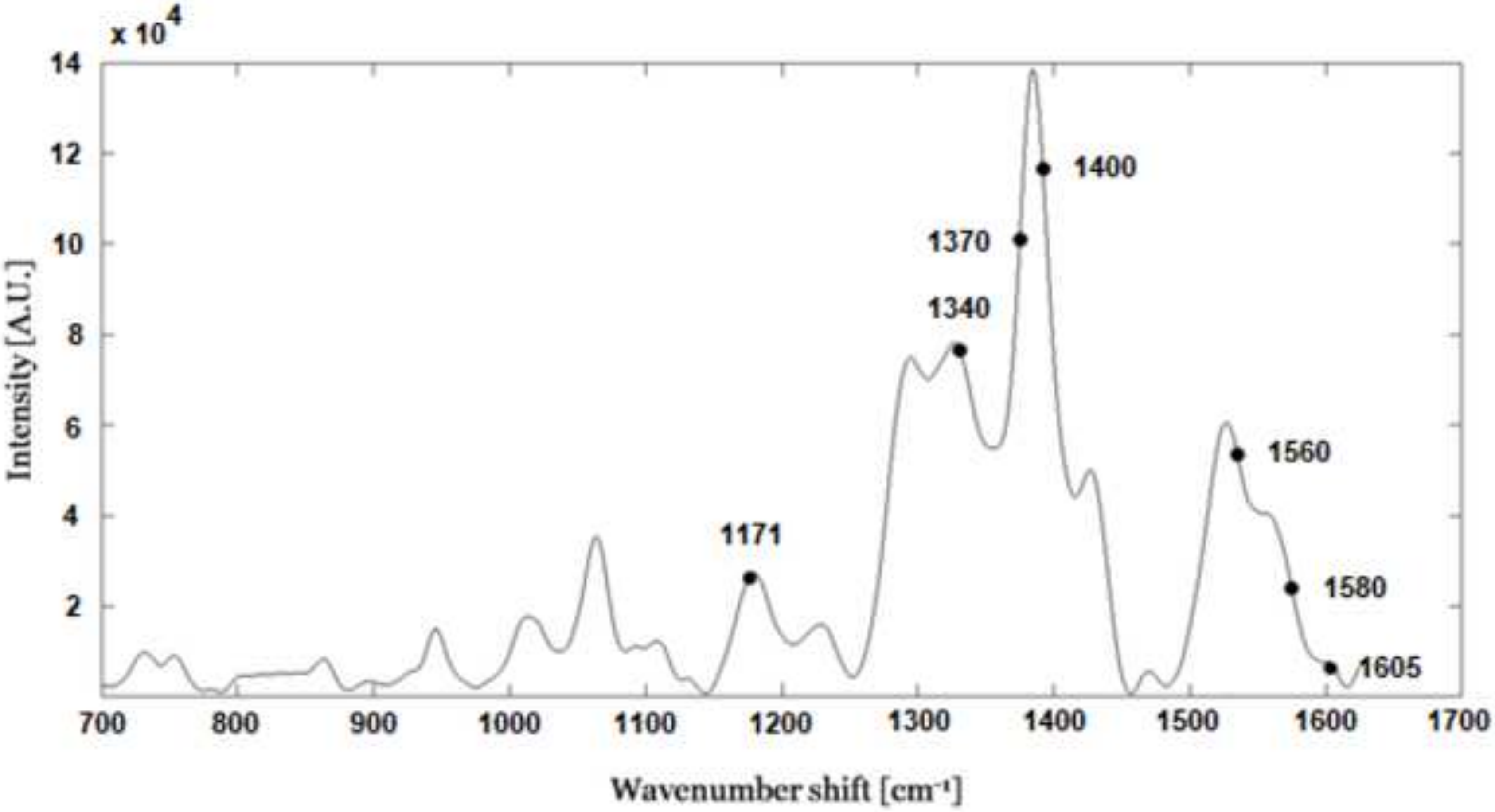
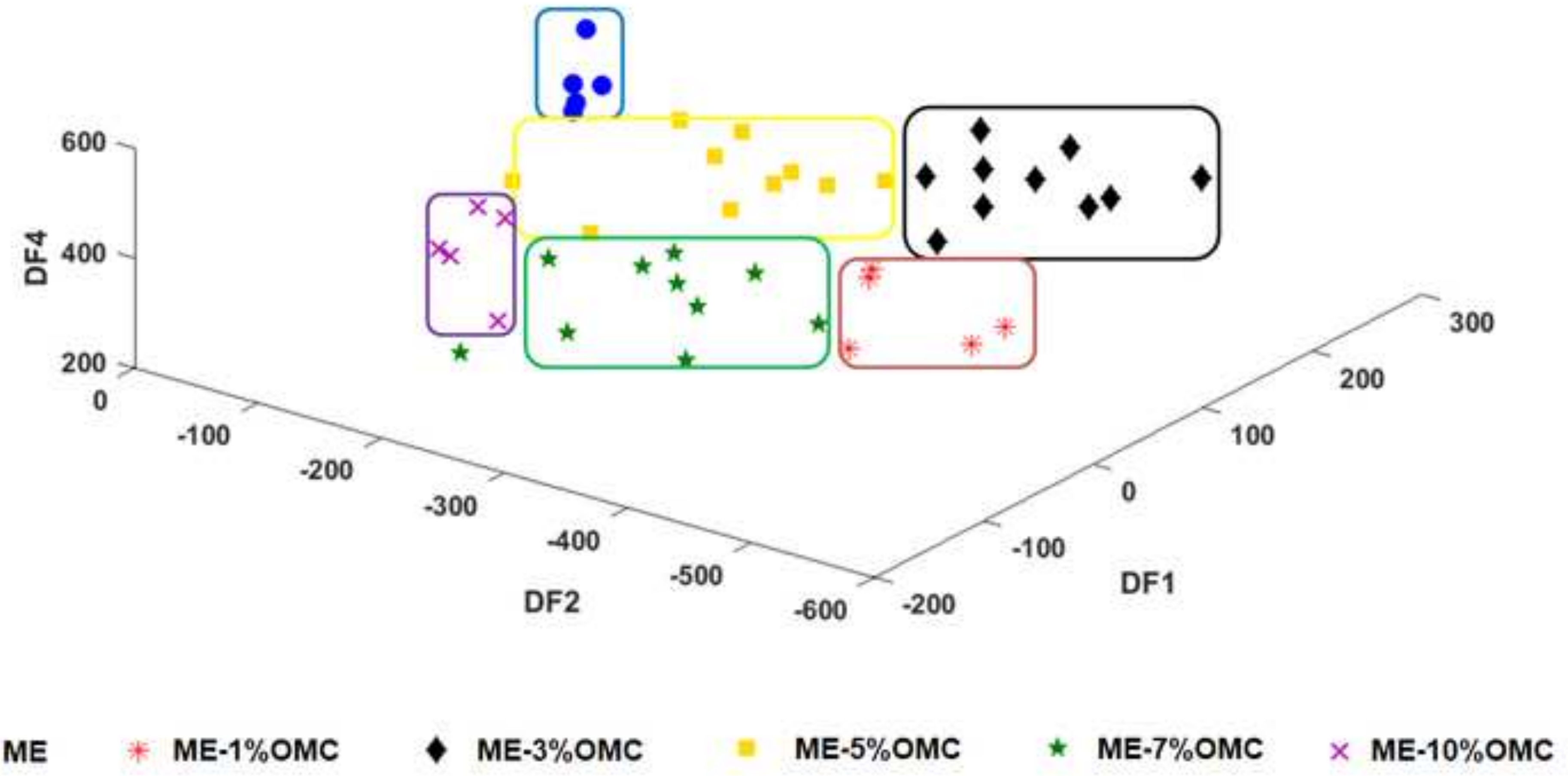
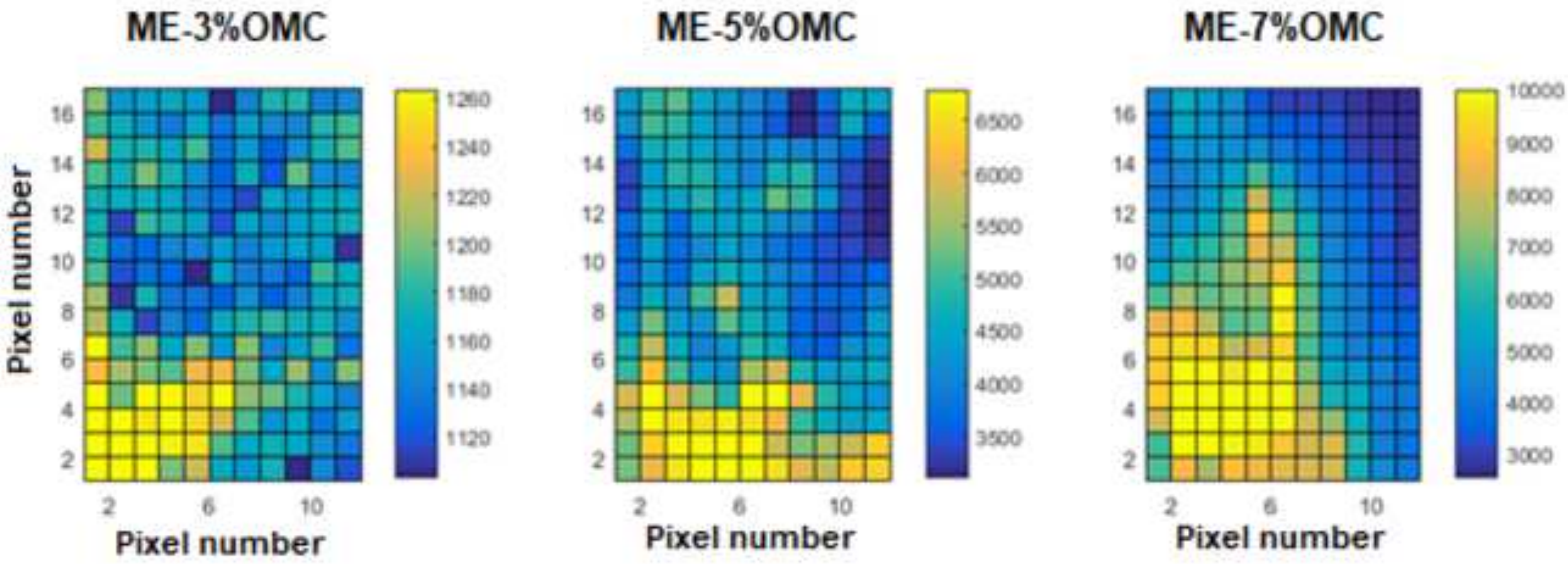

x 10^4
14
12
10
8
6
4
2
Intensity [A.U.]
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
Wavenumber shift [cm⁻¹]
1171
1340
1370
1400
1560
1580
1605

Figure 3


DF4
600
400
200
0
-100
-200
-300
-400
-500
-600
DF2
-200
-100
0
100
200
300
DF1
● ME
✳ ME-1%OMC
◆ ME-3%OMC
■ ME-5%OMC
★ ME-7%OMC
× ME-10%OMC


ME-3%OMC
Pixel number
16
14
12
10
8
6
4
2
2
6
10
Pixel number
1260
1240
1220
1200
1180
1160
1140
1120
ME-5%OMC
Pixel number
6500
6000
5500
5000
4500
4000
3500
ME-7%OMC
Pixel number
10000
9000
8000
7000
6000
5000
4000
3000

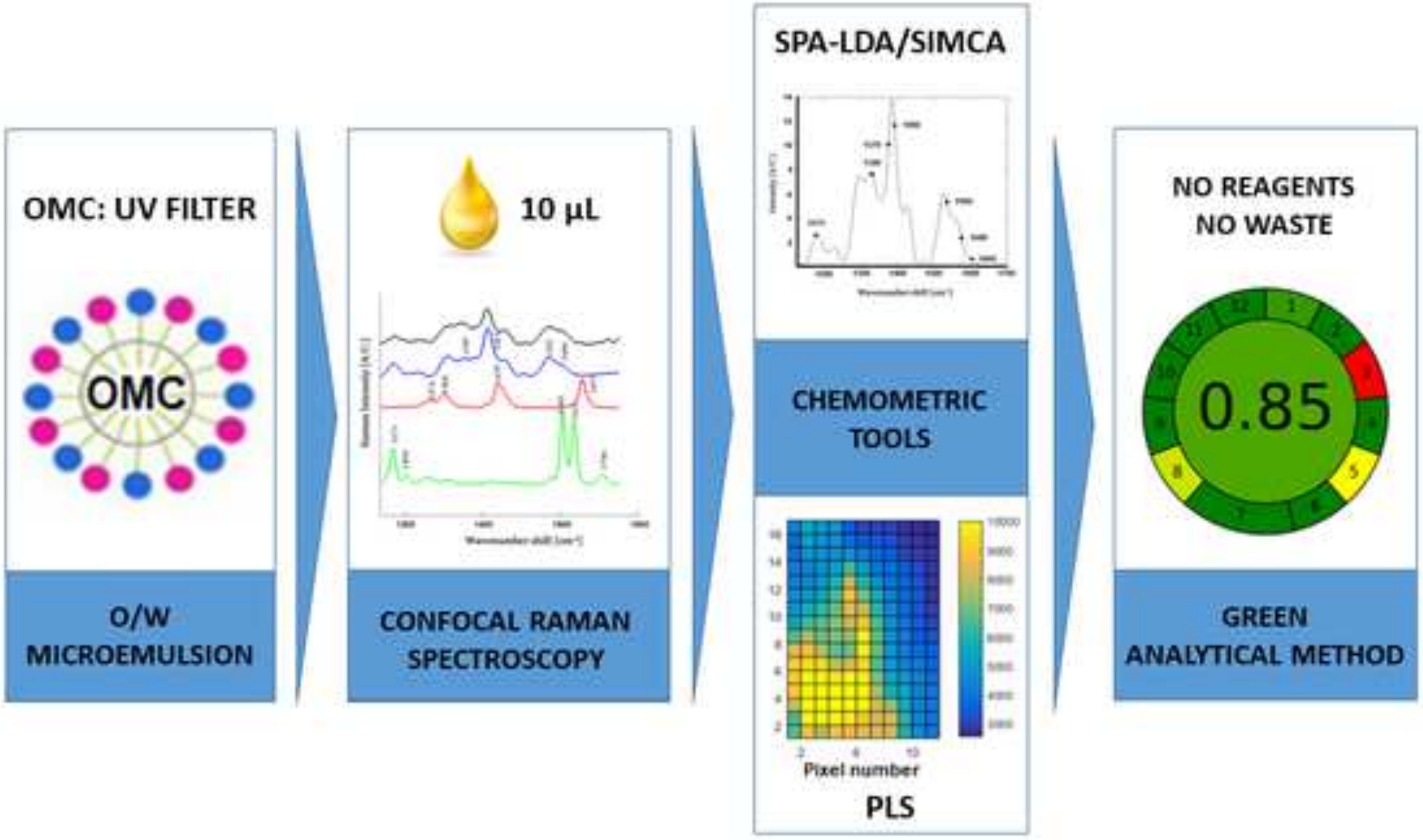
Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

AUTHOR STATEMENT

Danielle Silva do Nascimento: Formal Analysis, Investigation, Writing - Original Draft. Verónica Volpe: Validation. Cintia Fernandez: Formal Analysis. Martín Oresti: Validation. Lorna Ashton: Resources, Writing - Review & Editing, Supervision. Marcos Grünhut: Conceptualization, Writing - Review & Editing, Supervision.


OMC: UV FILTER
OMC
O/W
MICROEMULSION
10 µL
CONFOCAL RAMAN
SPECTROSCOPY
SPA-LDA/SIMCA
CHEMOMETRIC
TOOLS
16
14
12
10
8
6
4
2
2
6
10
Pixel number
PLS
NO REAGENTS
NO WASTE
0.85
GREEN
ANALYTICAL METHOD

Confocal Raman spectroscopy assisted by chemometric tools: A green approach for classification and quantification of octyl p-methoxycinnamate in oil-in-water microemulsions

Danielle Silva do Nascimento[a], *Verónica Volpe*[a], *Cintia Fernandez*[a], *Martín Oresti*[b], *Lorna Ashton*[c], *Marcos Grünhut*[a*]

[a] INQUISUR (UNS-CONICET), Department of Chemistry, Universidad Nacional del Sur, Bahía Blanca, Argentina.

[b] INIBIBB (UNS-CONICET), Departamento de Biología, Bioquímica y Farmacia, Universidad Nacional del Sur, Bahía Blanca, Argentina

[c] Department of Chemistry, Lancaster University, Lancaster, United Kingdom

***Corresponding author e-mail:**

E-mail: mgrunhut@uns.edu.ar

Abstract

This work proposes a green analytical method based on confocal Raman spectrometry and chemometrics tools for the qualitative and quantitative analysis of oil in water microemulsions loaded with the UVB filter octyl p-methoxycinnamate (OMC). The method does not use reagents and only 10 µL of sample are needed. The analyzed microemulsion samples were synthetized in the laboratory using decaethylene glycol mono-dodecyl ether (21.9%) as non-ionic surfactant, ethyl alcohol (7.3%) as co-surfactant, oleic acid (1.5%) as oil phase and water (69.3%). A physicochemical characterization of the samples was carried out obtaining expected values for droplet size (<20 nm), polydispersity index (<0.290) and conductivity (0.04–0.07 mS cm^{-1}), among others. Linear discriminant analysis (LDA) after selection of variables using the successive projections algorithm (SPA) and soft independent modelling of class analogy (SIMCA) were employed to classify microemulsions with different concentrations of OMC (1.0 to 10.0%). In the case of LDA, seven Raman spectral variables were previously selected by SPA and after this SPA-LDA model resulted in one error in the prediction set achieving an accuracy of 98%. The SIMCA model (α=0.05) presented an explained variance higher 97% using four principal components and it allowed the correct classification of 100% of the samples (N=15). In the quantitative analysis, partial least squares (PLS) was used to determine OMC in a range according to international legislation. The model presented optimal statistical parameters (R^2=0.9699; RMSEP=0.54%) and the prediction of samples were in close agreement with HPLC method. Moreover, the greenery of the method was estimated using the AGREE metric and an optimal value of 0.85 was obtained demonstrating the proposed analytical method results environmentally friendly.

Keywords: UV filters; *O*ctyl p-methoxycinnamate; Microemulsions; Confocal Raman spectroscopy; Linear Discriminant Analysis; Successive Projections Algorithm; Soft Independent Modelling by Class Analogy; Partial Least Squares; Green Analytical Chemistry

1. Introduction

Pharmaceutical and personal care products (PPCPs) comprise different kind of formulations that include products for sun protection, which present one or more chemical species that can reduce the negative effects of solar radiation [1,2]. Chemical sunscreens are organic molecules with chromophoric groups that absorb radiation in the UV region [3,4]. Octyl p-methoxycinnamate (OMC) is a lipophilic UVB filter (290-320 nm) belonging to the methoxycinnamate group widely used in PPCPs [5]. However, some side effects are associated with OMC such as minor skin irritation and dryness or hardening of the skin. Rarely, people may develop more serious skin side effects such as burning, stinging, and swelling or rash [6]. On the other hand, concentrations of OMC lower than declared in the product represents a significant risk to the health since the protection against sun exposure would not be guaranteed. Thus, OMC is usually present in the concentration range from 0.1% to 10% (w/w) in PPCPs taking into account the technical specifications of the Food and Drug Administration (US) and the European Union, which authorize a maximum level of 7.5% (w/w) and 10% (w/w), respectively [7,8]. On this form, the quality control of formulation containing OMC is necessary.

It is important to note that OMC decreases its efficacy as a sunscreen when exposed to sunlight. The direct photolysis of the OMC produces their photoisomerization, changing octyl p-methoxy-trans-cinnamate (trans-OMC) to octyl p-methoxy-cis-cinnamate (cis-OMC). The molar absorption coefficient of the cis-isomer (ε_{cis} = 12,600 L mol^{-1} cm^{-1}) is markedly lower than trans-isomer (ε_{trans} = 24,000 L mol^{-1} cm^{-1}) resulting in a reduction of the OMC efficiency as UV filter [9]. This problem can be solved using appropriate vehicles as carriers of OMC such as microemulsion systems [10]. Microemulsions (MEs) are organized, transparent and thermodynamically stable systems widely used in PPCPs for the topical administration of various components. These systems consist of two immiscible liquids (aqueous phase and oily phase) stabilized by the presence of a surfactant and, in some cases, a cosurfactant [11]. MEs present advantages such as the ability to protect labile compounds increasing the bioavailability. Besides, because of the presence of both lipophilic and hydrophilic domains, MEs are appropriate systems to incorporate a wide range of molecules [12]. However, the analytical study of this kind of systems may become difficult due to their physicochemical characteristics and the high concentration of components with different solubility (surfactant, co-surfactant, oily phase).

The methods applied for the determination of OMC in PPCPs are based on chromatographic techniques such as high-performance liquid chromatography (HPLC) [13] and gas chromatography (GC) [14], amperometric techniques [15], and spectroscopic techniques such as UV-Vis spectroscopy [16], mass spectrometry [17], nuclear magnetic resonance spectroscopy [18], infrared spectroscopy [19] and Raman spectroscopy [20]. Since OMC is usually present in complex matrices, most of these spectroscopic techniques cannot be used for direct analysis without sample pretreatment [21,22]. Confocal Raman spectroscopy (CRS) has been widely used to determine the composition and structure of different materials. The main advantages of CRS are allowing the analysis of samples without previous pretreatment; a rapid analysis; and it is a non-destructive technique, among others [23,24]. The use of CRS combined with chemometrics tools is presented as a powerful analytical alternative since relevant information can be extracted from the Raman spectra to obtain classification and quantification models for different species in complex samples [25-29]. Soft independent modeling of class analogy (SIMCA) is a supervised pattern recognition method that uses principal component analysis (PCA) to model the hyperspace of each class [30]. The PCA method promotes compression of a large data set and the variance is concentrated in few variables called principal components. However, when employing a large data set in the construction of mathematic models, such as a full spectrum, many variables are redundant and/or non-informative, and their inclusion may affect the quality of the final model. To overcome this drawback, the use of variable selection techniques is an appropiate alternative [31]. Successive projection algorithm (SPA) is the forward algorithm, with restriction that the selected variable in each interation is the least collinear to other selected variables [32]. Pontes et al. [33] adapted SPA so that it can be used with linear discriminant analysis (LDA) in the variable selection to solve classification problems. SPA-LDA is employed as supervised classification technique that focuses at selecting a subset of variables with minimal collinearity and appropriate discriminating ability [34]. On the other hand, partial least squares (PLS) is a well known algorithm applied successfully for building regression models. PLS technique makes use of the inverse calibration approach, where it is possible to calibrate for the desired component while implicitly modeling the other sources of variation. The inversion problem is resolved by replacing the original variables with linear combinations of the variables normally called factors. The optimum number of factors is often chosen on the basis of the quality of the results for the samples in the prediction set [30].

The combination of CRS with classification and calibration techniques is well documented for a wide type of samples. However, the use of CRS assisted by chemometrics tools in MEs systems is still limited [35].

As is known, the analytical determination of drugs is normally carried out by separation techniques, such as HPLC and GC. These techniques use organic solvents that are not environmentally friendly and are normally energy and time-consuming. The principles of green analytical chemistry (GAC) emphasize the use of safer and less toxic solvents, as well as reducing energy consumption and waste generation, avoiding derivatization, and favoring the use of substances based on renewable sources, among others [36]. The correct application of GAC provides many benefits in various aspects of sustainability. GAC comprises 12 principles that serve as a guide to defining the framework of green analytical procedures. In the last years, a trend to design environmentally friendly procedures in analytical chemistry has been highlighted [37,38] and the adoption of this kind of procedures is mandatory for adherence to ISO14000 guidelines. However, the application of green chemistry metrics for the evaluation of the greenery of analytical methods is restricted [39].

This work proposes classification and multivariate calibration green methods for the analysis of OMC loaded MEs samples applying chemometrics tools to the spectral data obtained by CRS. Class discrimination was carried out for samples containing different concentrations of OMC (1.0 to 10%) using the SIMCA and SPA-LDA algorithms. Moreover, the quantification of OMC in MEs was carried out using PLS algorithm. Moreover, the degree of the greenness was estimated by using the green metric analytical approach (AGREE).

2. Material and Methods

2.1 Reagents

All the reagents were analytical-grade chemicals and ultra-pure water (18 MΩ cm^{-1}) was obtained from a Barnsted® water purification system. Decaethylene glycol mono-dodecyl ether (DME; Sigma-Aldrich®) and ethyl alcohol (ET; Dorwil®) were used as non-ionic surfactant and co-surfactant, respectively. Oleic acid (OA; Applichem®) was used as oil phase and octyl-p-methoxy cinnamate (OMC; Parafarm®) was used as UVB filter.

2.2 Preparation of the OMC loaded microemulsions

An o/w microemulsion based on biocompatible materials and composed of 21.9% of DME (surfactant), 7.3% of ET (cosurfactant), 1.5% of OA (oil phase) and 69.3% of water was used for this study and obtained as described by the authors in a previous work [9]. Briefly, different amounts of OMC were added to the oil phase by stirring at 25 °C in order to obtain a final UV filter concentration of 1.0, 3.0, 5.0, 7.0 and 10.0% (w/w). These dosages were chosen taking into account the technical specifications for the United States Food and Drugs Administration and the European Union, which authorize a maximum level of 7.5 and 10% (w/w) of OMC, respectively [8,9]. Also, it is well known that microemulsions used as delivery vehicles of lipophilic molecules should kept both surfactant and co-surfactant levels as low as possible. In this case, the used concentrations of both DME (21.9%) and ET (7.3%) were appropriate for this kind of formulations. Then, the mixture OMC-oil phase was added to the DME:ET (3:1) mixture and titrated with ultra-pure water under moderate magnetic stirring. All prepared OMC loaded MEs were visually examined for transparency and stored in amber-glass containers at room temperature (25 °C).

2.3 Characterization of the OMC loaded microemulsions

Different parameters were studied in order to corroborate the structure, stability and capability of the obtained o/w microemulsions.

Average droplet size (Z) and polydispersity index (PdI) were determined by dynamic light scattering (DLS) using a Malvern Zetasizer Nano Series instrument (Malvern, UK). After 10-fold dilution in ultra-pure water, the samples were magnetically stirred for 5 min and the measurements were carried out at 25 °C, in an optical quality 4 mL borosilicate cell and at 90° angle. Then, Z and PdI were calculated by the Zetasizer 7.13 software (Malvern, UK) using the time correlation function. The conductivity of the preparations (25 °C) was determined using an Adwa model AD32 conductivity meter (Adwa, Hungary). The pH measurements were performed by using an Orion model 710 A pH meter with an Orion-Ross® model 81-02 electrode (Thermo, USA). For this, 1.0 g of sample was dispersed in 10 mL of ultra-pure water and stirred during 60 min at 25 °C.

In addition, the sun protection factor (SPF) was determinate by the in vitro Mansur method [40]. Briefly, the o/w microemulsions were diluted in ethanol at a final concentration of 0.2 mL mL^{-1}. Then, the spectrophotometric measurements between 290 and 320 nm with an interval of 5 nm were obtained by using a Hewlet Packard 8453 UV-Vis spectrophotometer (Agilent Technologies, Inc., USA) equipped with a quartz cell (10 mm optical path). The SPF was calculated as follows:

$$\mathbf{SPF} = \mathbf{CF} \sum_{290}^{320} \mathbf{EE}_{(\lambda)}\ \mathbf{I}_{(\lambda)}\ \mathbf{A}_{(\lambda)} \qquad \text{Eq. 1}$$

where CF is a correction factor (in this case, 10), EE is the erythemal effect spectrum, I is the solar intensity spectrum and A is the measured absorbance of the OMC-ME at each wavelength (λ). The relationship between the erythemal effect spectrum and the solar intensity spectrum at each wavelength (EE×I) was determined as described by Sayre et al. [41] (Table S1).

2.4 Spectral data acquisition

The Raman spectra were acquired with an inVia™ confocal microscope (Renishaw, UK) equipped with a 532 nm excitation wavelength laser and a L50x objective. The Wire 4.1 software (Renishaw, UK) was used to acquire the Raman spectra in the range from 600 to 1700 cm^{-1}. 10 µL of sample were prepared on glass slides and the scattered light was collected in the backscattering geometry onto a 2400 l/mm grating and a cooled charged-coupled device (CCD). All spectra were single accumulation collected for 20 s. Five spectra per sample were collected and then randomized. The samples were scanned in a streamline HR mode with 0.8 s exposure and 100% laser power. Each sample generated is aimed to produce quantitative concentration maps constituted from 17 x 12 pixels of area for each formulation.

2.5 Preparation of training and prediction sets

The Kennard-Stone (KS) uniform sampling algorithm [42] was used for the construction of both training and prediction sets. The training set contained samples in six levels of OMC concentrations ranging from 0 to 10% (0, 1.0, 3.0, 5.0, 7.0 and 10% w/w). In the case of the prediction set, it presented samples with three levels of OMC concentrations (3.0, 5.0 and 7.0% w/w). Five replicates of each sample were prepared rendering a total of 30 and 15 samples of training and prediction sets, respectively (Table S2). Both sets were used to perform the classification (LDA and SIMCA) and quantitative (PLS) analysis.

2.6 Data analysis

2.6.1 *Spectra preprocessing and exploratory analysis*. Usually, Raman spectra require a baseline correction. Asymmetric least squares (AsLS) algorithm [43] is adequate since it avoids the semi manual and individual correction of baseline that is subjective and time consuming. The parameters asymmetry (p) and smoothness (λ) were tested and the selected values were 0.001 and 10^3, respectively. In addition, the spectra were smoothed using the Savitzky-Golay (SG) algorithm [44] with a second-order polynomial and a 7-point window. On the other hand, PCA algorithm was used as unsupervised pattern recognition technique in the exploratory analysis of the spectral data.

2.6.2 Multivariate classification models. SPA-LDA and SIMCA were employed as supervised pattern recognition techniques for screening analysis.

LDA is focused on finding optimal boundaries between classes, maximizing between-class separability while minimizing within-class variability. However, LDA requires a selection of variables when full-spectrum data is available [45]. In SPA-LDA a training set consisting of n samples with known class labels is used to perform the process of variable selection. In the case of Raman spectroscopy, each sample consists of a spectrum with k wavenumbers. At first, the n training samples are centered on the mean of each class and stacked in the form of a **X** matrix ($n \times k$), where each column of **X** corresponds to a variable. Then, projection operations related to the columns of **X** are carried out to create k chains with l variables. Due to the loss of freedom degrees in the process of calculating class means, the chain length is limited by $n - c$, where c is the number of classes involved in the problem. Each time, the chain is initialized by one of the available k variables. Subsequent variables are selected to the chain in order to display the least collinearity with the previous ones, which is evaluated by the correlation between the respective column vectors of **X**. Finally, a total of $k \times l$ subsets of variables can be generated. These candidate subsets are assessed in terms of a cost function involving the average risk of misclassification over the validation set [33]. On this form, the optimal number of variables (NVs) is established for SPA-LDA model.

SIMCA is based on the advantage of soft modeling where each class can be independently modeled. This classification technique is based on producing a PCA model for each class of sample and then comparing their distance to the class confidence limits. On this form, SIMCA allows the analysis of a large data set concentrating the variance in few variables called principal components (PCs). Considering a set of n objects (samples) measured in k sensors (wavenumbers in Raman spectroscopy) generating the **X** matrix ($n \times k$), PCA reduced the matrix **X** into a product of two other arrays of low dimensionality **T** (n x A) and **L** (k x A) called scores and loadings, respectively, where A represents the number of new variables considered to be significant for the model of each class. The new variables in **T** present the advantage of being mutually orthogonal allowing the use of all spectral information in the construction of the SIMCA model. This characteristic permits the detection of anomalous samples or outliers, present in the data set. Finally, the classification of new samples is carried out by means of an F-test at a given significance level. The optimal number of PC's can be selected by leave-one-out cross-validation using the training set by means of the calculation of the root mean square error of cross validation (RMSECV) [30].

The optimal NVs and PCs were established for SPA-LDA and SIMCA, respectively. The performance of the classification models was evaluated by accuracy, which is defined using the ratio of samples in the test set correctly assigned into their respective classes.

2.6.3 Multivariate calibration model. Regression analysis was performed by applying the PLS algorithm. PLS is a factor analysis method which the full-spectrum advantage is retained by forming a new coordinate system, as follow:

$$\mathbf{X} = \mathbf{T} \times \mathbf{L} + \mathbf{E} \qquad \text{Eq. 2}$$

where **X** is the m x n matrix of calibration spectra, **L** is a h x n matrix with the rows of **L** being the new PLS basis set of h full-spectrum vectors, often called loading vectors or loading spectra. **T** is an m x h matrix of intensities (scores) in the new coordinate system of the h PLS loading vectors for the m sample spectra. **E** is now the $m \times n$ matrix of spectral residuals not fit by the optimal PLS model. The spectral intensities in the new coordinate system can be related to concentrations using least-squares analysis:

$$\mathbf{c} = \mathbf{T} \times \mathbf{b} + \mathbf{e} \qquad \text{Eq. 3}$$

where **b** is the $h \times l$ vector of coefficients relating the scores to the concentrations, **T** is the matrix of scores from the PLS spectral decomposition (Eq. 2) and **e** an error vector. The concentration in an unknown sample is obtained by using the previously calculated **b** coefficients [30].

In this work, the PLS model and its performance was evaluated by leaving one out cross validation, in which each sample of the calibration set was left out once, and its concentration was estimated by a model built with the remaining samples. The models were evaluated by calculating analytical parameters as the determination coefficient (R^2), root mean square error of prediction (RMSEP) and the relative error of prediction (REP%). In addition, figures of merit as sensitivity (SEN) was calculated as SEN = $1/||b_k||$ and limit of detection (LOD) and limit of quantification (LOQ) were calculated as LOD (or LOQ) = $k\ \alpha\ ||b_k||$, where k= 3.3 for LOD, and k= 10 for LOQ, α is the standard deviation of the net analytical signal (NAS), and $||b_k||$ is the Euclidean norm of the vector of regression coefficients estimated from the multivariate calibration model for analyte k.

Multivariate data analysis including KS, AsLS, SG, SPA-LDA, SIMCA and PLS algorithms was performed using MatLab R2018 software (The MathWorks, USA).

2.7 HPLC analysis

The quantification of OMC in the o/w microemulsions was performed by using a HPLC Agilent 1260 Infinity system equipped with an Eclipse Plus (Agilent Technologies) C18 reverse phase column (3.5 μm spherical particles covered with octadecylsilane, 100 × 4.6 mm). Methanol and ultrapure water in an 85:15 (v/v) ratio was the mobile phase and was controlled by the activation of quaternary pumps at 25 °C. The flow rate was 0.5 mL min^{-1}. OMC was monitored at 311 nm using a UV detector. Moreover, isocratic mode was employed and the injected sample volume was 20 μL. An external calibration curve was constructed using standard solutions of OMC in the range of 10-150 mg L^{-1}. Finally, the peak area obtained for a retention time of 16.5 min was used for the OMC quantification.

3. Results and discussion

3.1 Characterization of the OMC loaded microemulsions

As shown in Table 1, the Z values for all formulations were less than 100 nm, varying between 9.99 and 15.14 nm, in agreement with expected values for microemulsions [12]. In addition, the size was increasing as the concentration of OMC increased. The corresponding histograms are showing in Fig. S1. Thus, from the size distribution it is possible to obtain the PdI, which is the ratio between the standard deviation and the mean droplet size. PdI values between 0.1 and 0.25 indicate a narrow size distribution whereas a PdI>0.5 indicates a very broad size distribution. All the formulations presented a PdI between 0.082 and 0.286 (Table 1), indicating that the microemulsions tended to approach monodispersion and become stable systems due to the uniformity in the drops.

On the other hand, the conductivity of the microemulsions were in range 0.04-0.07 mS cm^{-1}, demonstrating that all the systems were o/w type. Furthermore, as expected, the values of conductivity decreased as the OMC concentration was higher (Table 1). All the OMC-MEs presented a pH range between 3.5 and 4.0 which is suitable for topical use [11]. Finally, as expected, the SPF values were increasing as the OMC concentration in the formulation increased (Table 1).

3.2 Band assignments

The Raman spectra of OMC, DME, OA and the o/w microemulsion are shown in Fig. 1 along with the chemical structures of the microemulsion components. As can be seen, the spectrum of OMC (green line) shows five strong peaks around 1171, 1204, 1605, 1635 and 1706 cm^{-1}. The peak at 1171 cm^{-1} is assigned to the C–H bend, while the mode at 1204 cm^{-1} correspond to the C–O–C bend of the *p*-methoxy group. The modes at 1605 cm^{-1} are assigned to aromatic ring vibrations. In addition, the peaks at 1706 and 1635 cm^{-1} are assigned to the conjugated ester C=O stretch and the conjugated exocyclic C=C stretch, respectively [46]. In the case of the oleic acid (red line) the spectrum shows peaks at 1655, 1439 and 1302 cm^{-1}. The peak at 1655 cm^{-1} can be attributed to the stretching C=O bonds in the carboxylic group. The peak at 1439 cm^{-1} is assigned to CH_2 scissoring mode while the peak at 1302 cm^{-1} is attributable to CH_2 wagging and the carboxylic C–O stretching mode. In addition, the peaks at 1274, and 972 cm^{-1} can be assigned to stretching C–O bonds, and OH out of plane bending modes, respectively [47]. Concerning the DME spectrum (blue line), it shows the main peaks around 1606, 1562, 1413, 1347, and 1040 cm^{-1}. The peaks at 1606, 1562 and 1413 cm^{-1} are attributed to the stretching C=C bonds while the O-H in plane bending appears at 1347 cm^{-1}. Further, the band at 1040 cm^{-1} correspond to the C–O stretching vibration [48]. Finally, the ME spectrum (black line) was similar to the surfactant spectrum with bands in 1606, 1562, 1413, 1347 and 1040 cm^{-1} due to the fact that the ME contains DME as principal component (21.9%).

3.3 Multivariate analysis

Before the multivariate analysis, the spectra data were preprocessed as described in section 2.6.1. Fig. S2a and S2b show the Raman spectra corresponding to the training and the prediction sets before and after the data preprocessing step using AsLS and SG algorithms, respectively. Moreover, an exploratory analysis of the data was performed using PCA. As can be seen in Fig. S3, the graph of scores (PC1 vs PC2) shows that discrimination of the samples was not achieved and the use of supervised pattern recognition techniques was required.

3.3.1 Multivariate classification models

The statistic results corresponding to SPA-LDA and SIMCA methods obtained for the classification of o/w microemulsion samples containing different concentrations of OMC are presented in Table 2.

In the case of SPA-LDA, the minimum value of the cost function (0.378) was obtained using seven spectral variables. Fig. 2 shows a Raman spectrum of an OMC loaded microemulsion (5.0%) indicating the seven selected variables by SPA: 1171, 1340, 1370, 1400, 1560, 1580 and 1605 cm^{-1}. As can be seen, several variables agree with the regions of higher intensity for OMC (1171, 1580 and 1605 cm^{-1}). However, some variables outside these regions are important and have also been identified by SPA. For example, the variables 1340, 1370, 1400 and 1560 corresponding to the surfactant (DME) were also selected. Then, a LDA model was obtained and the graph of the Fisher discriminant functions (DF) are showed in Fig. 3. As can be seen, the discrimination between ME/ME-1%OMC and ME-1%OMC/ME-3%OMC samples in DF1 is observed, whereas DF2 discriminated ME-7%OMC/ME-10%OMC samples. Likewise, ME-3%OMC samples were discriminated from ME-5%OMC samples and ME-5%OMC samples were discriminated from ME-7%OMC samples in DF4. In terms of performance, the SPA-LDA model achieved an accuracy of 97.8% since only one sample corresponding to ME-7%OMC class was misclassified.

On the other hand, for SIMCA model the optimal number of principal components for defining each class was four with an explained variance higher than 97% (α=0.05). As can be seen in the Coomans graph (Fig. S4), the model allowed the correct classification of all analyzed samples achieving an accuracy of 100%. However, it is important to note that SIMCA used a total of 930 spectral variables in contrast to SPA-LDA which used only 7, reducing in a 99.2% the number of variables used to the construction of the classification model.

3.3.2 Multivariate calibration model

PLS regression was performed using six latent variables (PCs) which explained 97.5% of the total variance. The analytical parameters (R^2, RMSEP and REP%) and figures of merit (SEN, LOD and LOQ) of the PLS calibration model are shown in the Table 3. The R^2 (0.9699) was appropriate and both RMSEP (0.5434% (w/w)) and REP (9.88%) values were reasonably low indicating that the proposed method is accurate in the prediction of samples. This fact is important since that the samples contain high concentrations of other components such as oil (OA), surfactant (DME) and cosurfactant (ET) which could interfere significantly. With regard to the figures of merit, satisfactory values for SEN (52.19% (w/w)), LOD (0.0300% (w/w)) and LOQ (0.0990% (w/w)) were obtained considering the OMC concentration present in the analyzed formulations. However, the LOD and LOQ values were higher than those obtained by other authors using separation techniques as HPLC [49-53].

3.4 Mapping analysis

The ultra-high resolution (both spectrally and spatially) of the Raman microscope provides information about the identity and presence of specific components and reveals their location and distribution within a sample area [29]. The proposed PLS-CRS method allowed obtaining distribution maps (2D image) which provided a visual representation of the OMC distribution in the analyzed samples. A rainbow color-coded composite image was developed to show the localization of OMC. Fig. 4 shows the color spatial mapping corresponding to the predicted distribution of OMC in formulations with different concentration of OMC (3.0, 5.0 and 7.0%). The pixel concentration was predicted by multiplying each Raman spectrum by the corresponding PLS regression vector. As can be seen, over an area of 204 μm^2, the pixels that present orange-yellow color intensity correspond to higher concentrations of OMC, showing in light blue-blue regions with lower concentrations of OMC. As expected, the orange-yellow color pattern represented a larger area as the concentration of OMC increased. In relation to distribution, the OMC seems non-homogeneously distributed in the microemulsion samples. It is possible that oil drops have different drug loading capacity, which is reflected in regions with different concentrations of OMC. It is important to note that CRS was appropriate to predict the content of OMC and to depict its distribution in samples containing different concentrations of the UV-filter.

3.5 Validation of CRS-PLS method

Validation of trueness was carried out by comparison of the results obtained by the proposed method with the ones obtained by the reference method (HPLC). The statistical comparison was carried out using the joint interval test for the slope and the intercept [54]. The values obtained by the proposed method were regressed against the HPLC method. The estimated intercept (a) and slope (b) obtained, were compared with the ideal values of intercept = 0 and slope = 1. The elliptical joint confidence regions obtained (Fig. S5) indicate that there are no significant statistical differences between the results obtained by both methods, considering an overall significance level of $\alpha = 0.05$. From these results it can be concluded that, in comparison to the reference method, there is no bias in the results obtained for OMC.

On the other hand, Table 4 compares the proposed method with respect to different HPLC methods [49-53] developed to quantify OMC in different PPCPs. As can be seen, the proposed CRS-PLS method does not require a separation technique or the use of organic solvents. Moreover, the analysis time is short mainly due to the fact that a rapid instrumental response is obtained using CRS in comparison to the HPLC methods.

3.6 Green approach

In order to assess the greenness of an analytical methodology, it is necessary to have metrics to evaluate and easily quantify this parameter. The AGREE metric [55] is a recently proposed method that provides an analytical greenness estimation calculator, based on the twelve principles of GAC. The Fig. S6a and S6b show the colored pictogram corresponding to the proposed green CRS and HPLC methods, respectively. As can be seen, a numerical value of 0.85 was obtained for the proposed method, which is close to the ideal value of 1.0, demonstrating the greenery of the method. This fact is relating to a non-destructive analysis of the sample, requiring a minimal volume of it for the complete analysis (10 μL). In addition, there was no need of any pretreatment or derivatization steps, which significantly reduces the total time of analysis. The procedure did not include reagents and organic solvents and no waste was generated, which is an important benefit in terms of safety for both, the operator and the environment. Since the waste generation is one of the most important points in the evaluation of greenness of an analytical method we assigned a higher weight value to this parameter. This fact can be observed in the pictogram in which all the points represent the same area with the exception of number seven which correspond to waste. On the other hand, the points 3, 5 and 8 were not green, indicating a certain distance from the ideal value that is 1.0. As can be seen in Fig. S7, the point 3 is related to: *If possible, the measurements should be performed in situ*. In our case, this point was the farthest from the ideal greenness value, since the measurement device was in off-line position. Although these measurements were carried out off-line smaller at-line Raman spectrometers could be used instead once analytical methods are established. Point 5 states: *Automated and miniaturized method should be selected*. In the pictogram this point was yellow representing a smaller distance from the ideal value since the proposed method did not include automation. Finally, point 8 is related to: *Multi-analyte or multi-parameter methods are preferred versus methods using one analyte at a time*. This point was distant from the ideal value for two reasons: firstly, because only one analyte was determined and after because the throughput was 15 h^{-1}. This value was relatively less because included the chemometric analysis step. On the contrary, the pictogram corresponding to the HPLC method shows that the overall value of greenness provided by the metric is much lower (0.47). This fact is mainly due to the use of organic solvents required by the chromatographic technique, the absence of automation and/or miniaturization, and the impossibility of performing in situ analysis, among others. In particular, the use of methanol negatively impacts the analysis since this solvent is not obtained from renewable sources and it is a potentially dangerous solvent for the operator and the environment (toxic and highly flammable). Moreover, the HPLC method generated waste which should be avoided (Fig. S7). In conclusion, the value obtained for the proposed CRS method is much higher than the HPLC method and it corresponds to an excellent greenness.

4. Conclusions

This work demonstrated that the combination of confocal Raman spectroscopy and chemometrics is a potential tool to both qualitative and quantitative analysis of OMC loaded in oil in water microemulsions, alternatively to the traditionally chromatographic methods. In addition, confocal Raman spectroscopy allowed direct analysis of samples containing high concentration of surfactant, cosurfactant and oil phase, avoiding pretreatment steps, and obtaining fast, simple and reliable analytical methods in concordance whit the principles of green analytical chemistry. The combination of confocal Raman spectroscopy and SPA-LDA and SIMCA provided accurate methods for discriminate different percentages of OMC in microemulsion samples demonstrating the potential use of this methodology as a screening tool. On the other hand, a PLS model based on the analysis of Raman spectra exhibited optimal analytical parameters for the quantification of OMC at different concentration levels which represents an important fact since the concentration of OMC is directly related to the solar protection factor and therefore with the efficacy of the formulation. In addition, the proposed PLS-CRS method allowed obtaining distribution maps which provided a visual representation of the OMC distribution in the analyzed samples. Finally, the use of green chemistry metrics allowed to demonstrate the greenery of the proposed method obtaining a high value, which was close to the corresponding an ideal green methodology.

Acknowledgments

The authors gratefully acknowledge the financial support of Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET; PIP 11220120100625), Ministerio de Ciencia y Tecnología (PICT, 201-0659) and Universidad Nacional del Sur (PGI 24Q086 and PGI 24/B272).

References

[1] S. Dey, F. Bano, A. Malik, Pharmaceuticals and personal care products (PPCP) contamination- a global discharge inventory, in: Butterworth-Heinemann (Eds.), Phamarceuticals and personal care products: Waste Management and Treatment Technology, Oxford, 2019, pp. 1-26.

[2] S.E. Mancebo, J.Y. Hu, S.Q. Wang, Sunscreens: A review of health benefits, regulations and controversies, *Dermatol. Clin.* 32 (2014) 427-438. http://dx.doi.org/10.1016/j.det.2014.03.011

[3] R. Jansen, U. Osterwalder, S.Q. Wang, M. Burnett, H.W. Lim, Photoprotection: part II. Sunscreen: development, efficacy, and controversies, J. Am. Acad. Dermatol. 69 (2013) 867.e1-867.e14. https://doi.org/10.1016/j.jaad.2013.08.022

[4] J.C. Cardoso, B.M.L. Armondes, J.B.G. Júnior, T.A. de Araújo, J.L.R. Júnior, N.R. Poppi, V.S. Ferreira, Determination of 4-methylbenzilidene camphor in sunscreen by square wave voltammetry in media of cationic surfactant, Microchem. J. 85 (2007) 301-307. https://doi.org/10.1016/j.microc.2006.07.006

[5] D.S. Nascimento, V. Volpe, M. Insausti, M. Grünhut, On line kinetic analysis of permeation profiles for UV filter loaded microemulsions using an automatic system with spectroscopic detection and a chemometric approach, RSC Adv. 11 (2021) 15528-15538. http://doi.org/10.1039/D1RA01067K

[6] G.J. Nohynek, H. Schaefer, Benefit and risk of organic ultraviolet filters, *Regul. Toxicol. Phamarcol.* 33 (2001) 285-299. http://doi.org/10.1006/rtph.2001.1476

[7] FDA, Department of Health and Human Services, 21 CFR Parts 310, 352, 700 and 740, RIN 0910-AA01, Sunscreen Drug Products for the Over-the-Counter Human Use Final Monograph, Federal Register, Rules and Regulations, 64 (1999) 27666-27693. https://www.govinfo.gov/content/pkg/FR-1999-05-21/pdf/99-12853.pdf

[8] EU, European Directive 76/768/EEC and its successive amendments, basic act 31976L0768. https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=celex:31976L0768

[9] D.S. Nascimento, M. Insausti, B.S.F. Band, M. Grünhut, Photolysis study of octyl p-methoxycinnamate loaded microemulsion by molecular fluorescence and chemometric approach, Spectrochim. Acta A. 191 (2018) 277-282. http://dx.doi.org/10.1016/j.saa.2017.10.025

[10] V. Volpe, D.S. Nascimento, M. Insausti, M. Grünhut, Octyl p-methoxycinnamate loaded microemulsion based on Ocimum basilicum essential oil. Characterization and analytical studies for potential cosmetic applications, Colloids Surf. A Physicochem. Eng. Asp. 546 (2018) 285-292. https://doi.org/10.1016/j.colsurfa.2018.02.070

[11] M. Fanun, Microemulsions as delivery systems, Curr. Opin. Colloid Interface Sci. 17 (2012) 306-313. https://doi.org/10.1016/j.cocis.2012.06.001

[12] V. Volpe, M.N. Giacomodonato, D.O. Sordelli, M. Insausti, F.R. Buzzola, M. Grünhut, Ciprofloxacin loaded o/w microemulsion against Staphylococcus aureus. Analytical and biological studies for topical and intranasal administration, J. Drug Deliv. Sci. Technol. 57 (2020) 101705. https://doi.org/10.1016/j.jddst.2020.101705

[13] S. Scalia, Determination of sunscreen agents in cosmetic products by supercritical fluid extraction and high-performance liquid chromatography, J. Chromatogr. A. 870 (2000) 199-205. http://doi.org/10.1016/s0021-9673(99)01224-8

[14] L. Rubio, L. Valverde-Som, L.A. Sarabia, M.C. Ortiz, Improvement in the identification and quantification of UV filters and additives in sunscreen cosmetic creams by gas chromatography/mass spectrometry through three-way calibration

techniques, Talanta. 205 (2019) 120156. https://doi.org/10.1016/j.talanta.2019.120156

[15] V.S. Ferreira, J.B.G. Júnior, C.M.S.C. Oliveira, R.M. Takeuchi, A.L. Santos, M.A.G. Trindade, Voltammetric analysis of sun-block preparations containing octocrylene and its association with 2-Hydroxy-4-methoxybenzophenone and octyl-methoxycinnamate, Microchem. J. 106 (2013) 378-383. https://doi.org/10.1016/j.microc.2012.10.002.

[16] A. Chisvert, M.T. Vidal, A. Salvador, Sequential injection analysis for benzophenone-4 and phenylbenzimidazole sulphonic acid in sunscreen sprays by solid-phase extraction coupled with ultraviolet spectrometry, Anal. Chim. Acta. 464 (2002) 295-301. https://doi.org/10.1016/S0003-2670(02)00432-4

[17] M. Haunschmidt, W. Buchberger, C.W. Klampfl, R. Hertsens, Identification and semi-quantitative analysis of parabens and UV filters in cosmetic products by direct-analysis-in-real-time mass spectrometry and gas chromatography with mass spectrometric detection, Anal. Methods. 3 (2011) 99-104. https://doi.org/10.1039/C0AY00588F

[18] K. Mori, K. Itoh, S. Suzuki, H. Nakamura, Analysis of ultraviolet absorbers in cosmetics by two dimension NMR spectroscopy, J. Toxicol. Environ. Health. 42 (1996) 60-66. http://doi.org/10.1248/jhs1956.42.60

[19] X.H. Liu, X.Z. Liang, X. Fang, W.P. Zhang, Preparation and evaluation of novel octylmethoxycinnamate- loaded solid lipid nanoparticles, Int. J. Cosmet. Sci. 37 (2015) 446-453. https://doi.org/10.1111/ics.12216

[20] S.A. Oladepo, G.R. Loppnow, Ultraviolet resonance Raman spectroscopy as a robust spectroscopic tool for in situ sunscreen analysis Anal. Chim. Acta. 628 (2008) 57-66. http://doi.org/10.1016/j.aca.2008.08.031

[21] M. Mei, X. Huang, Online analysis of five organic ultraviolet filters in environmental water samples using magnetism-enhanced monolith-based in-tube

solid phase microextraction coupled with high-performance liquid chromatography, J. Chromatogr. A. 1525 (2017) 1-9. https://doi.org/10.1016/j.chroma.2017.09.065

[22] M. Vila, J.P. Lamas, C. Garcia-Jares, T. Dagnac, M. Llompart, Ultrasound-assisted emulsification microextraction followed by gas chromatography-mass spectrometry and gas chromatography-tandem mass spectrometry for the analysis of UV filters in water, Microchem. J. 124 (2016) 530-539. https://doi.org/10.1016/j.microc.2015.09.023

[23] F.D. Fleischli, S. Mathes, C. Adlhart, Label free non-invasive imaging of topically applied actives in reconstructed human epidermis by confocal Raman spectroscopy, Vib. Spectrosc. 68 (2013) 29-33. http://doi.org/10.1016/j.vibspec.2013.05.003

[24] M. Sohn, T. Buehler, G. Imanidi, Repartition of oil miscible and water soluble UV filters in an applied sunscreen film determined by confocal Raman microspectroscopy, Photochem Photobiol Sci. 15 (2016) 861-871. https://doi.org/10.1039/C6PP00024J

[25] M.C. Breitkreitz, G.P. Sabin, G. Polla, R.J. Poppi, Characterization of semi-solid self-emulsifying drug delivery systems (SEDDS) of atorvastatin calcium by Raman image spectroscopy and chemometrics, J. Pharm. Biomed. Anal. 73 (2013) 3-12. http://doi.org/10.1016/j.jpba.2012.03.054

[26] E.M. Paiva, V.H. Silva, R.J. Poppi, C.F. Pereira, J.J.H. Rohwedder, Comparison of macro and micro Raman measurement for reliable quantitative analysis of pharmaceutical polymorphs J. Pharm. Biomed. Anal. 157 (2018) 107-115. http://doi.org/10.1016/j.jpba.2018.05.010

[27] D.J. Silva, H. Wiebeck, Using PLS, iPLS and siPLS linear regressions to determine the composition of LDPE/HDPE blends: A comparation between confocal Raman and ATR-FTIR spectrocopies, Vib. Spectrosc. 92 (2017) 259-266. https://doi.org/10.1016/j.vibspec.2017.08.009

[28] L.M. Duarte, D. Paschoal, C.M.S. Izumi, M.D. Dolzan, V.R. Alves, G.A. Micke, H.F. Santos, M.A.L. Oliveira, Simultaneous determination of aspartame, cyclamate,

saccharin and acesulfame-K in powder tabletop sweeteners by FT-Raman spectroscopy associated with the multivariate calibration: PLS, iPLS and siPLS models were compared, Int. Food Res. J. 99 (2017) 106-114. https://doi.org/10.1016/j.foodres.2017.05.006

[29] A.A. Saleh, M. Hegazy, S. Abbas, A. Elkosasy, Development of distribution maps of spectrally similar degradation products by Raman chemical imaging microscope coupled with a new variable selection technique and SIMCA classifier, Spectrochim. Acta A Mol. Biomol. Spectrosc. 268 (2022) 120654. https://doi.org/10.1016/j.saa.2021.120654

[30] K.R. Beebe, R.J. Pell, M.B. Seasholtz, Chemometrics: A Practical Guide. John Wiley & Sons, inc. 1998.

[31] C.M. Andersen, R. Bro, Variable selection in regression-A Tutorial, J. Chemom. 24 (2010) 728-737. https://doi.org/10.1002/cem.1360

[32] M.C.U. Araújo, T.C.B. Saldanha, R.K.H Galvão, T. Yoneyama, H.C. Chame, V. Visani, The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. Chemom. Intell. Lab. Syst. 57 (2001) 57-65. https://doi.org/10.1016/S0169

[33] M.J.C. Pontes, R.K.H. Galvão, M.C.U. Araujo, P.N.T. Moreira, O.D. Pessoa Neto, G.E. José, T.C.B. Saldanha, The successive projections algorithm for spectral variable selection in classification problems, Chemom. Intell. Lab. Syst. 78 (2005) 11-18. https://doi.org/10.1016/j.chemolab.2004.12.001

[34] C.C. Acebal, M. Grünhut, A.G. Lista, B.S.F. Band, Successive projections algorithm applied to spectral data for the simultaneous determination of flavour enhancers, Talanta 82 (2010) 222-226. https://doi.org/10.1016/j.talanta.2010.04.024

[35] K. Wang, D.W. Sun, Q. Wei, H. Pu, Quantification and visualization of α-tocopherol in oil-in-water emulsion based delivery systems by Raman microspectroscopy, LTW-Food Sci. Technol. 96 (2018) 66-74. https://doi.org/10.1016/j.lwt.2018.05.017

[36] P.T. Anastas, M.M. Kirchhoff, Origins, current status, and future challenges of green chemistry. Acc. Chem. Res. 35 (2002) 686-694. https://doi.org/10.1021/ar010065m

[37] M.F. Razuc, M. Grünhut, E. Saidman, M. Garrido, B.S.F. Band, Green method based on a flow-batch analyzer system for the simultaneous determination of ciprofloxacin and dexamethasone in pharmaceuticals using a chemometric approach, Talanta. 115 (2013) 314-322. https://doi.org/10.1016/j.talanta.2013.05.002

[38] M. Espino, F.J.V. Gómez, J. Boiteux, M.Á. Fernández, M.F. Silva, Green Chemistry Metrics, Comprehensive Foodomics. 2021, pp. 825-833. https://doi.org/10.1016/B978-0-08-100596-5.22822-9

[39] A. Gałuszka, Z.M. Migaszewski, P. Konieczka, J. Namiésnik, Analytical Eco-Scale for assessing the greenness of analytical procedures, TrAC Trends Analyt. Chem. 37, (2012) 61-72. https://doi.org/10.1016/j.trac.2012.03.013

[40] J.S. Mansur, M.N. Breder, M.C. Mansur, R.D. Azulay, Determination of sun protection factor by spectrophotometry, An. Bras. Dermatol. 61 (1986) 121-124.

[41] R.M. Sayre, P.P. Agin, G.J. Levee, E. Marlowe, A comparison of in vivo and in vitro testing of sunscreening formulas, Photochem. Photobiol. 29 (1979) 559-566. http://dx.doi.org/10.1111/j.1751-1097.1979.tb07090.x

[42] R.W. Kennard, L.A. Stone, Computer-aided design of experiments. Technometrics, 11 (1969) 137-148. https://doi.org/10.2307/1266770

[43] H. Shixuan, W. Zhang, L. Liu, L. Huang, J. He, W. Xie, P. Wu, C. Du, Baseline Correction for Raman Spectra Using Improved Asymmetric Least Squares, Anal. Methods. 6 (2014) 4402-4407. https://doi.org/10.1039/C4AY00068D

[44] A. Savitzky, M.J.E. Golay, Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Anal. Chem. 36 (1964) 1627-1639. https://pubs.acs.org/doi/pdf/10.1021/ac60214a047

[45] D.D. de Sousa Fernandes, V.E. de Almeida, M.M. Fontes, M.C.U. de Araújo, G. Véras, P.H.G.D. Diniz, Simultaneous identification of the wood types in aged cachaças

and their adulterations with wood extracts using digital images and SPA-LDA, Food Chem. 273 (2018) 77-84. https://doi.org/10.1016/j.foodchem.2018.02.035

[46] L. Beyere, S. Yarasi and G.R. Loppnow, Solvent effects on sunscreen active ingredients using Raman spectroscopy, J. Raman Spectrosc. 34 (2003) 743-750. https://doi.org/10.1002/jrs.1042

[47] W.N. Silverio, G.V.S. Dutra, A.T. Jensen, O.A. Araújo, V. Garg, A.C. Oliveira, L.F. Valadares, F.G. Souza, F. Machado, Superparamagnetic nanoparticles stabilized with free-radical polymerizable oleic acid-based coating, J. Alloy Compd. 739 (2018) 1025-1036. https://doi.org/10.1016/j.jallcom.2017.12.338

[48] S. Sun, R.L. Birke, J.R. Lombardi, Surface-Enhanced Raman Spectroscopy of Surfactants on Silver Electrodes, J. Phys. Chem. 94 (1990) 2005-2010. https://doi.org/10.1021/j100368a052

[49] L. Montenegro, C. Carbone, G. Puglisi, Vehicle effects on in vitro release and skin permeation of octylmethoxycinnamate from microemulsions, Int. J. Pharm. 405 (2011) 162-168. https://doi.org/10.1016/j.ijpharm.2010.11.036

[50] E. Ramón, C. Alonso, L. Coderch, A. de la Maza, O. López, J.L. Parra, J. Notario, Liposomes as Alternative Vehicles for Sun Filter Formulations, Drug Deliv. 12 (2005) 83-88. https://doi.org/10.1080/10717540490446080

[51] M.S. Monteiro, R.A. Ozzetti, A.L. Vergnanini, L. de Brito-Gitirana, N.M. Volpato, Z.M. de Freitas, E. Ricci-Júnior, E.P. dos Santos, Evaluation of octyl p-methoxycinnamate included in liposomes and cyclodextrins in anti-solar preparations: preparations, characterizations and in vitro penetration studies, Int. J. Nanomedicine. 7 (2012) 3045-3058. http://doi.org/10.2147/IJN.S28550

[52] M. Wharton, M. Geary, N. O'Connor, B. Murphy, A rapid high performance liquid chromatographic (HPLC) method for the simultaneous determination of seven UV filters found in sunscreen and cosmetics, Int. J. Cosmet. Sci. 33 (2011) 164-170. https://doi.org/10.1111/j.1468-2494.2010.00607.x

[53] J. Teixeira, Z.M.F. de Freitas, M.S. de Monteiro, A.B. Vermelho, E. Ricci Junior, E.P. dos Santos, Development and characterization of photoprotective formulations containing keratin particles, Braz. J. Pharm. Sci. 55 (2019) 17649. https://doi.org/10.1590/s2175-97902019000217649

[54] N. Draper, H. Smith, Applied Regression Analysis, 2° Ed., John Wiley and Sons, New York, 1981.

[55] F. Pena-Pereira, W. Wojnowski, M.Tobiszewski, AGREE-Analytical GREEnness Metric Approach and Software, Anal. Chem. 92 (2020) 10076-10082. https://dx.doi.org/10.1021/acs.analchem.0c01887

Tables

Table 1. Physicochemical characterization of the analyzed OMC loaded o/w microemulsions (1.0; 3.0; 5.0; 7.0 and 10%). Z: average droplet size; PdI: polydispersity index; C: conductivity; SPF: solar protection factor. All measurements were performed in triplicate.

Sample	Z (nm)	PdI	C (mS cm^{-1})	pH	SPF
ME-1%OMC	9.99 ± 0.04	0.101 ± 0.070	0.070 ± 0.002	4.10 ± 0.01	1.73 ± 0.04
ME-3%OMC	9.93 ± 0.11	0.086 ± 0.045	0.060 ± 0.001	4.10 ± 0.01	3.92 ± 0.03
ME-5%OMC	10.25 ± 0.05	0.286 ± 0.046	0.050 ± 0.001	4.20 ± 0.03	4.12 ± 0.01
ME-7%OMC	13.71 ± 0.23	0.281 ± 0.071	0.050 ± 0.002	4.00 ± 0.01	5.13 ± 0.04
ME-10%OMC	15.14 ± 1.36	0.285 ± 0.018	0.040 ± 0.001	4.10 ± 0.02	18.20 ± 0.07

Table 2. Summary of the statistical results (95% confidence level) for SPA-LDA and SIMCA models applied to confocal Raman spectra data for the classification of o/w microemulsion samples containing different concentration of OMC.

Sample data set	SPA-LDA			SIMCA		
	NVs	Error	Accuracy (%)	PCs	Error	Accuracy (%)
Training (N=30)	7	0	100	4	0	100
Prediction (N=15)		1	97.8		0	100

NVs: number of variables; PCs: principal components

Table 3. Analytical parameters of the PLS calibration model for OMC determination in microemulsion samples.

Analytical parameter	Value
Spectral range (cm^{-1})	600-1700
Concentration range (% w/w)	0-10
Latent variables	6
R^2	0.9699
RMSEP (% w/w)	0.5434
REP (%)	9.88
SEN (% w/w)	52.19
LOD (% w/w)	0.0300
LOQ (% w/w)	0.0990

Table 4. Comparison of methods developed to determination of OMC in different PPCPs.

Method	Separation technique	Detection technique	Sample	Solvent	Time analysis (s)
[45]	HPLC	UVS (310 nm)	o/w microemulsion	ACN : Water (80:20; v/v)	ND
[46]	HPLC	UVS (307 nm)	o/w emulsion	Water : MeOH (12:88; v/v)	330
[47]	HPLC	UVS (310 nm)	Gel cream	Water : MeOH : (13:87; v/v)	ND
[48]	HPLC	UVS (313 nm)	Sunscreen/Cosmetics	EtOH : Acetic Acid 1% (70:30; v/v)	420
[49]	HPLC	UVS (320 nm)	Sunscreen	ACN:Water (90:10; v/v)	600
This work	Not used	CRS (600-1700 cm^{-1})	o/w microemulsion	Not used	20

UVS: UV spectroscopy; ND: not declared

Figure captions

Figure 1. Raman spectra corresponding to OMC-ME (black line), DME (blue line), OA (red line) and OMC (green line). Wavenumbers (cm^{-1}) of the principal peaks are indicated along with the chemical structure of the components. The spectra have been shifted along the y-axis for a better comprehension.

Figure 2. Raman spectrum corresponding to a OMC loaded microemulsion (5.0%). The circles in black indicates the seven variables selected by SPA algorithm (1171, 1340, 1370, 1400, 1560, 1580 and 1605 cm^{-1}).

Figure 3. Fisher's discriminant functions (DF) obtained by SPA-LDA applied to Raman data set. The samples are represented as ME (blue circle), ME-1%OMC (red asterisk), ME-3%OMC (black diamond), ME-5%OMC (yellow square), ME-7%OMC (green star) and ME-10%OMC (violet cross).

Figure 4. Distribution maps representing the OMC predicted concentration by PLS regression in different o/w microemulsion samples (ME-1.0%OMC; ME-3.0%OMC; ME-7.0%OMC). The pixels that present highest colour intensity (orange-yellow) correspond to higher concentrations of OMC.

Figure 1

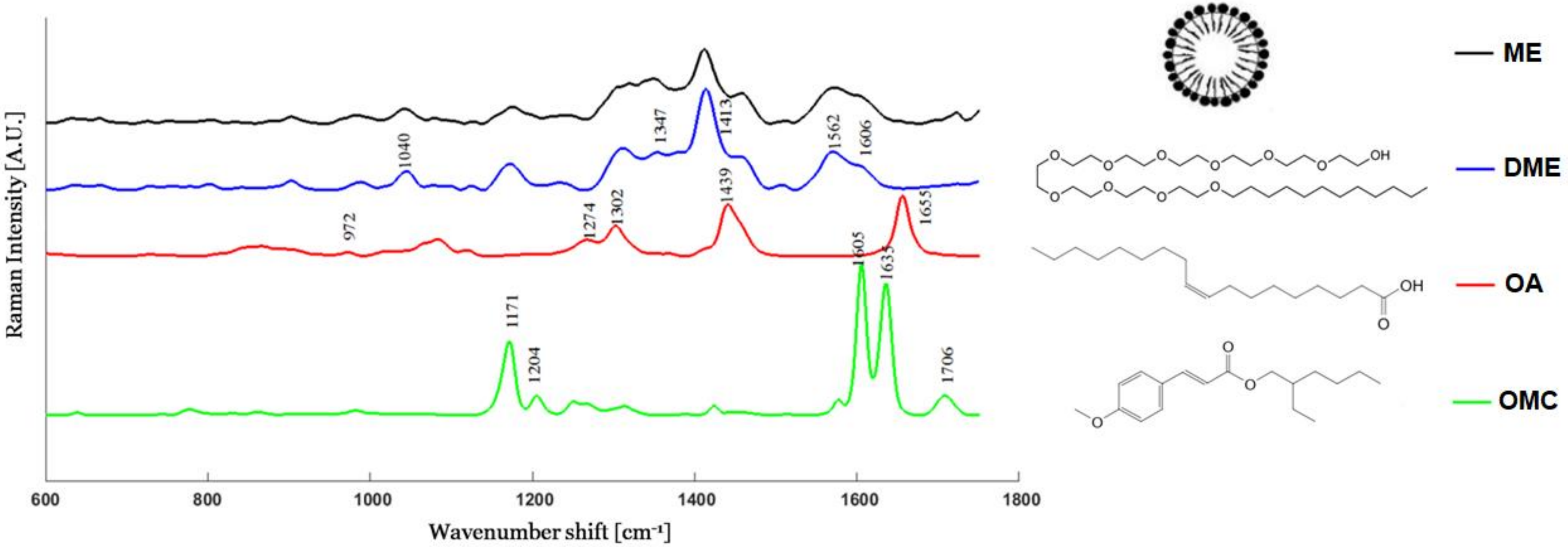
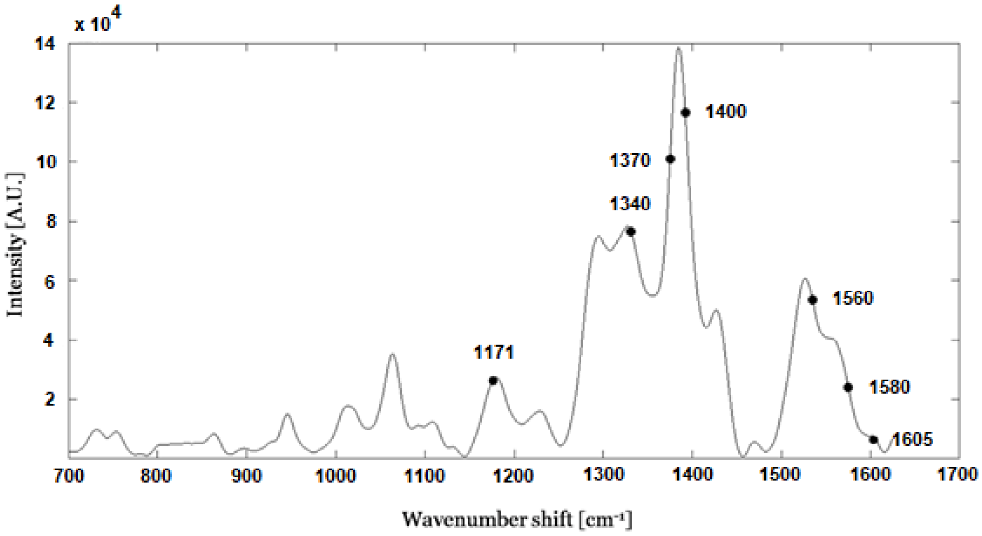

Raman Intensity [A.U.]
Wavenumber shift [cm⁻¹]
600
800
1000
1200
1400
1600
1800
1347
1413
1562
1606
1040
972
1274
1302
1439
1655
1605
1635
1171
1204
1706
O
OH
O
OH
O
O
ME
DME
OA
OMC

Figure 2



x 10^4
14
12
10
8
6
4
2
Intensity [A.U.]
1400
1370
1340
1560
1171
1580
1605
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
Wavenumber shift [cm⁻¹]

Figure 3

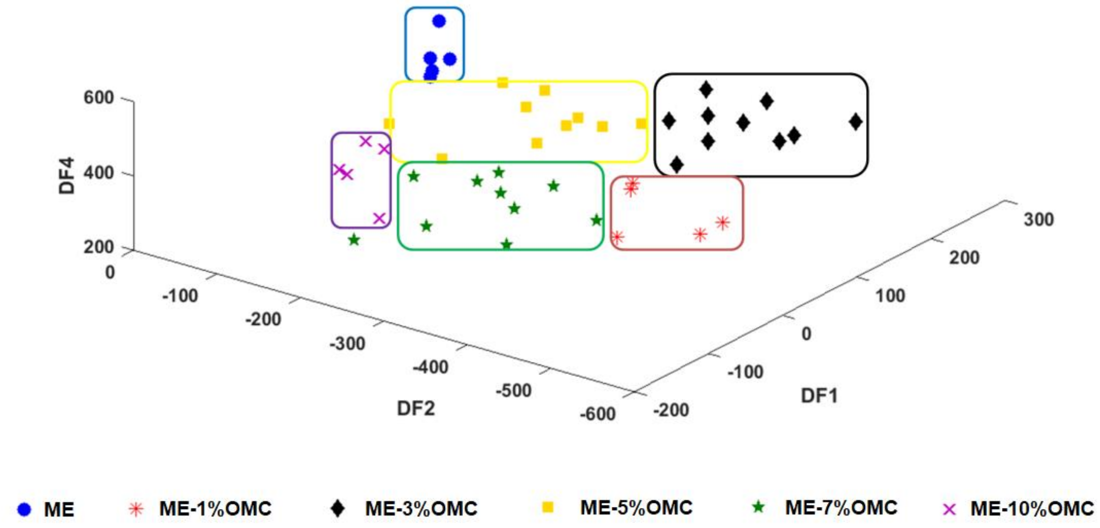
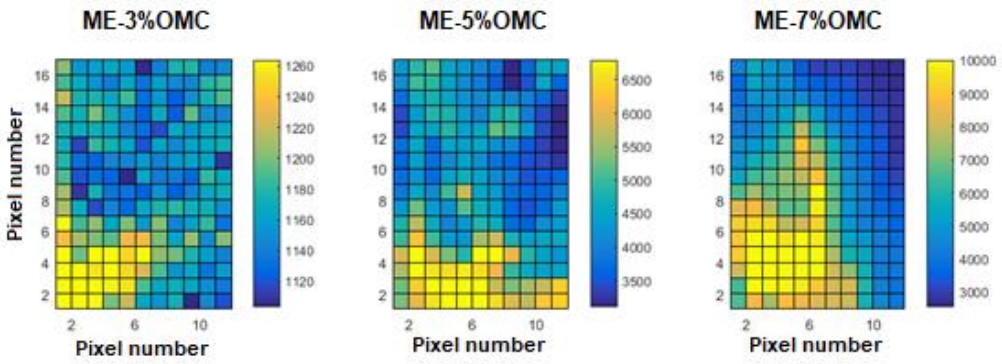

600
400
200
DF4
0
-100
-200
-300
-400
-500
-600
DF2
-200
-100
0
100
200
300
DF1
ME
ME-1%OMC
ME-3%OMC
ME-5%OMC
ME-7%OMC
ME-10%OMC

Figure 4


ME-3%OMC
Pixel number
16
14
12
10
8
6
4
2
2
6
10
Pixel number
1260
1240
1220
1200
1180
1160
1140
1120
ME-5%OMC
16
14
12
10
8
6
4
2
2
6
10
Pixel number
6500
6000
5500
5000
4500
4000
3500
ME-7%OMC
16
14
12
10
8
6
4
2
2
6
10
Pixel number
10000
9000
8000
7000
6000
5000
4000
3000

Confocal Raman spectroscopy assisted by chemometric tools: A green approach for classification and quantification of octyl p-methoxycinnamate in oil-in-water microemulsions

Danielle Silva do Nascimento[a], *Verónica Volpe*[a], *Cintia Fernandez*[a], *Martín Oresti*[b], *Lorna Ashton*[c], *Marcos Grünhut*[a*]

[a] INQUISUR (UNS-CONICET), Department of Chemistry, Universidad Nacional del Sur, Bahía Blanca, Argentina.

[b] INIBIBB (UNS-CONICET), Departamento de Biología, Bioquímica y Farmacia, Universidad Nacional del Sur, Bahía Blanca, Argentina

[c] Department of Chemistry, Lancaster University, Lancaster, United Kingdom

***Corresponding author e-mail:**

E-mail: mgrunhut@uns.edu.ar

Abstract

This work proposes a green analytical method based on confocal Raman spectrometry and chemometrics tools for the qualitative and quantitative analysis of oil in water microemulsions loaded with the UVB filter octyl p-methoxycinnamate (OMC). The method does not use reagents and only 10 µL of sample are needed. The analyzed microemulsion samples were synthetized in the laboratory using decaethylene glycol mono-dodecyl ether (21.9%) as non-ionic surfactant, ethyl alcohol (7.3%) as co-surfactant, oleic acid (1.5%) as oil phase and water (69.3%). A physicochemical characterization of the samples was carried out obtaining expected values for droplet size (<20 nm), polydispersity index (<0.290) and conductivity (0.04–0.07 mS cm^{-1}), among others. Linear discriminant analysis (LDA) after selection of variables using the successive projections algorithm (SPA) and soft independent modelling of class analogy (SIMCA) were employed to classify microemulsions with different concentrations of OMC (1.0 to 10.0%). In the case of LDA, seven Raman spectral variables were previously selected by SPA and after this SPA-LDA model resulted in one error in the prediction set achieving an accuracy of 98%. The SIMCA model (α=0.05) presented an explained variance higher 97% using four principal components and it allowed the correct classification of 100% of the samples (N=15). In the quantitative analysis, partial least squares (PLS) was used to determine OMC in a range according to international legislation. The model presented optimal statistical parameters (R^2=0.9699; RMSEP=0.54%) and the prediction of samples were in close agreement with HPLC method. Moreover, the greenery of the method was estimated using the AGREE metric and an optimal value of 0.85 was obtained demonstrating the proposed analytical method results environmentally friendly.

Keywords: UV filters; *O*ctyl p-methoxycinnamate; Microemulsions; Confocal Raman spectroscopy; Linear Discriminant Analysis; Successive Projections Algorithm; Soft Independent Modelling by Class Analogy; Partial Least Squares; Green Analytical Chemistry

1. Introduction

Pharmaceutical and personal care products (PPCPs) comprise different kind of formulations that include products for sun protection, which present one or more chemical species that can reduce the negative effects of solar radiation [1,2]. Chemical sunscreens are organic molecules with chromophoric groups that absorb radiation in the UV region [3,4]. Octyl p-methoxycinnamate (OMC) is a lipophilic UVB filter (290-320 nm) belonging to the methoxycinnamate group widely used in PPCPs [5]. However, some side effects are associated with OMC such as minor skin irritation and dryness or hardening of the skin. Rarely, people may develop more serious skin side effects such as burning, stinging, and swelling or rash [6]. On the other hand, concentrations of OMC lower than declared in the product represents a significant risk to the health since the protection against sun exposure would not be guaranteed. Thus, OMC is usually present in the concentration range from 0.1% to 10% (w/w) in PPCPs taking into account the technical specifications of the Food and Drug Administration (US) and the European Union, which authorize a maximum level of 7.5% (w/w) and 10% (w/w), respectively [7,8]. On this form, the quality control of formulation containing OMC is necessary.

It is important to note that OMC decreases its efficacy as a sunscreen when exposed to sunlight. The direct photolysis of the OMC produces their photoisomerization, changing octyl p-methoxy-trans-cinnamate (trans-OMC) to octyl p-methoxy-cis-cinnamate (cis-OMC). The molar absorption coefficient of the cis-isomer (ε_{cis} = 12,600 L mol^{-1} cm^{-1}) is markedly lower than trans-isomer (ε_{trans} = 24,000 L mol^{-1} cm^{-1}) resulting in a reduction of the OMC efficiency as UV filter [9]. This problem can be solved using appropriate vehicles as carriers of OMC such as microemulsion systems [10]. Microemulsions (MEs) are organized, transparent and thermodynamically stable systems widely used in PPCPs for the topical administration of various components. These systems consist of two immiscible liquids (aqueous phase and oily phase) stabilized by the presence of a surfactant and, in some cases, a cosurfactant [11]. MEs present advantages such as the ability to protect labile compounds increasing the bioavailability. Besides, because of the presence of both lipophilic and hydrophilic domains, MEs are appropriate systems to incorporate a wide range of molecules [12]. However, the analytical study of this kind of systems may become difficult due to their physicochemical characteristics and the high concentration of components with different solubility (surfactant, co-surfactant, oily phase).

The methods applied for the determination of OMC in PPCPs are based on chromatographic techniques such as high-performance liquid chromatography (HPLC) [13] and gas chromatography (GC) [14], amperometric techniques [15], and spectroscopic techniques such as UV-Vis spectroscopy [16], mass spectrometry [17], nuclear magnetic resonance spectroscopy [18], infrared spectroscopy [19] and Raman spectroscopy [20]. Since OMC is usually present in complex matrices, most of these spectroscopic techniques cannot be used for direct analysis without sample pretreatment [21,22]. Confocal Raman spectroscopy (CRS) has been widely used to determine the composition and structure of different materials. The main advantages of CRS are allowing the analysis of samples without previous pretreatment; a rapid analysis; and it is a non-destructive technique, among others [23,24]. The use of CRS combined with chemometrics tools is presented as a powerful analytical alternative since relevant information can be extracted from the Raman spectra to obtain classification and quantification models for different species in complex samples [25-29]. Soft independent modeling of class analogy (SIMCA) is a supervised pattern recognition method that uses principal component analysis (PCA) to model the hyperspace of each class [30]. The PCA method promotes compression of a large data set and the variance is concentrated in few variables called principal components. However, when employing a large data set in the construction of mathematic models, such as a full spectrum, many variables are redundant and/or non-informative, and their inclusion may affect the quality of the final model. To overcome this drawback, the use of variable selection techniques is an appropiate alternative [31]. Successive projection algorithm (SPA) is the forward algorithm, with restriction that the selected variable in each interation is the least collinear to other selected variables [32]. Pontes et al. [33] adapted SPA so that it can be used with linear discriminant analysis (LDA) in the variable selection to solve classification problems. SPA-LDA is employed as supervised classification technique that focuses at selecting a subset of variables with minimal collinearity and appropriate discriminating ability [34]. On the other hand, partial least squares (PLS) is a well known algorithm applied successfully for building regression models. PLS technique makes use of the inverse calibration approach, where it is possible to calibrate for the desired component while implicitly modeling the other sources of variation. The inversion problem is resolved by replacing the original variables with linear combinations of the variables normally called factors. The optimum number of factors is often chosen on the basis of the quality of the results for the samples in the prediction set [30].

The combination of CRS with classification and calibration techniques is well documented for a wide type of samples. However, the use of CRS assisted by chemometrics tools in MEs systems is still limited [35].

As is known, the analytical determination of drugs is normally carried out by separation techniques, such as HPLC and GC. These techniques use organic solvents that are not environmentally friendly and are normally energy and time-consuming. The principles of green analytical chemistry (GAC) emphasize the use of safer and less toxic solvents, as well as reducing energy consumption and waste generation, avoiding derivatization, and favoring the use of substances based on renewable sources, among others [36]. The correct application of GAC provides many benefits in various aspects of sustainability. GAC comprises 12 principles that serve as a guide to defining the framework of green analytical procedures. In the last years, a trend to design environmentally friendly procedures in analytical chemistry has been highlighted [37,38] and the adoption of this kind of procedures is mandatory for adherence to ISO14000 guidelines. However, the application of green chemistry metrics for the evaluation of the greenery of analytical methods is restricted [39].

This work proposes classification and multivariate calibration green methods for the analysis of OMC loaded MEs samples applying chemometrics tools to the spectral data obtained by CRS. Class discrimination was carried out for samples containing different concentrations of OMC (1.0 to 10%) using the SIMCA and SPA-LDA algorithms. Moreover, the quantification of OMC in MEs was carried out using PLS algorithm. Moreover, the degree of the greenness was estimated by using the green metric analytical approach (AGREE).

2. Material and Methods

2.1 Reagents

All the reagents were analytical-grade chemicals and ultra-pure water (18 MΩ cm^{-1}) was obtained from a Barnsted® water purification system. Decaethylene glycol mono-dodecyl ether (DME; Sigma-Aldrich®) and ethyl alcohol (ET; Dorwil®) were used as non-ionic surfactant and co-surfactant, respectively. Oleic acid (OA; Applichem®) was used as oil phase and octyl-p-methoxy cinnamate (OMC; Parafarm®) was used as UVB filter.

2.2 Preparation of the OMC loaded microemulsions

An o/w microemulsion based on biocompatible materials and composed of 21.9% of DME (surfactant), 7.3% of ET (cosurfactant), 1.5% of OA (oil phase) and 69.3% of water was used for this study and obtained as described by the authors in a previous work [9]. Briefly, different amounts of OMC were added to the oil phase by stirring at 25 °C in order to obtain a final UV filter concentration of 1.0, 3.0, 5.0, 7.0 and 10.0% (w/w). These dosages were chosen taking into account the technical specifications for the United States Food and Drugs Administration and the European Union, which authorize a maximum level of 7.5 and 10% (w/w) of OMC, respectively [8,9]. Also, it is well known that microemulsions used as delivery vehicles of lipophilic molecules should kept both surfactant and co-surfactant levels as low as possible. In this case, the used concentrations of both DME (21.9%) and ET (7.3%) were appropriate for this kind of formulations. Then, the mixture OMC-oil phase was added to the DME:ET (3:1) mixture and titrated with ultra-pure water under moderate magnetic stirring. All prepared OMC loaded MEs were visually examined for transparency and stored in amber-glass containers at room temperature (25 °C).

2.3 Characterization of the OMC loaded microemulsions

Different parameters were studied in order to corroborate the structure, stability and capability of the obtained o/w microemulsions.

Average droplet size (Z) and polydispersity index (PdI) were determined by dynamic light scattering (DLS) using a Malvern Zetasizer Nano Series instrument (Malvern, UK). After 10-fold dilution in ultra-pure water, the samples were magnetically stirred for 5 min and the measurements were carried out at 25 °C, in an optical quality 4 mL borosilicate cell and at 90° angle. Then, Z and PdI were calculated by the Zetasizer 7.13 software (Malvern, UK) using the time correlation function. The conductivity of the preparations (25 °C) was determined using an Adwa model AD32 conductivity meter (Adwa, Hungary). The pH measurements were performed by using an Orion model 710 A pH meter with an Orion-Ross® model 81-02 electrode (Thermo, USA). For this, 1.0 g of sample was dispersed in 10 mL of ultra-pure water and stirred during 60 min at 25 °C.

In addition, the sun protection factor (SPF) was determinate by the in vitro Mansur method [40]. Briefly, the o/w microemulsions were diluted in ethanol at a final concentration of 0.2 mL mL^{-1}. Then, the spectrophotometric measurements between 290 and 320 nm with an interval of 5 nm were obtained by using a Hewlet Packard 8453 UV-Vis spectrophotometer (Agilent Technologies, Inc., USA) equipped with a quartz cell (10 mm optical path). The SPF was calculated as follows:

$$\mathbf{SPF} = \mathbf{CF} \sum_{290}^{320} \mathbf{EE}_{(\lambda)}\ \mathbf{I}_{(\lambda)}\ \mathbf{A}_{(\lambda)} \qquad \text{Eq. 1}$$

where CF is a correction factor (in this case, 10), EE is the erythemal effect spectrum, I is the solar intensity spectrum and A is the measured absorbance of the OMC-ME at each wavelength (λ). The relationship between the erythemal effect spectrum and the solar intensity spectrum at each wavelength (EE×I) was determined as described by Sayre et al. [41] (Table S1).

2.4 Spectral data acquisition

The Raman spectra were acquired with an inVia™ confocal microscope (Renishaw, UK) equipped with a 532 nm excitation wavelength laser and a L50x objective. The Wire 4.1 software (Renishaw, UK) was used to acquire the Raman spectra in the range from 600 to 1700 cm^{-1}. 10 µL of sample were prepared on glass slides and the scattered light was collected in the backscattering geometry onto a 2400 l/mm grating and a cooled charged-coupled device (CCD). All spectra were single accumulation collected for 20 s. Five spectra per sample were collected and then randomized. The samples were scanned in a streamline HR mode with 0.8 s exposure and 100% laser power. Each sample generated is aimed to produce quantitative concentration maps constituted from 17 x 12 pixels of area for each formulation.

2.5 Preparation of training and prediction sets

The Kennard-Stone (KS) uniform sampling algorithm [42] was used for the construction of both training and prediction sets. The training set contained samples in six levels of OMC concentrations ranging from 0 to 10% (0, 1.0, 3.0, 5.0, 7.0 and 10% w/w). In the case of the prediction set, it presented samples with three levels of OMC concentrations (3.0, 5.0 and 7.0% w/w). Five replicates of each sample were prepared rendering a total of 30 and 15 samples of training and prediction sets, respectively (Table S2). Both sets were used to perform the classification (LDA and SIMCA) and quantitative (PLS) analysis.

2.6 Data analysis

2.6.1 *Spectra preprocessing and exploratory analysis*. Usually, Raman spectra require a baseline correction. Asymmetric least squares (AsLS) algorithm [43] is adequate since it avoids the semi manual and individual correction of baseline that is subjective and time consuming. The parameters asymmetry (p) and smoothness (λ) were tested and the selected values were 0.001 and 10^3, respectively. In addition, the spectra were smoothed using the Savitzky-Golay (SG) algorithm [44] with a second-order polynomial and a 7-point window. On the other hand, PCA algorithm was used as unsupervised pattern recognition technique in the exploratory analysis of the spectral data.

2.6.2 Multivariate classification models. SPA-LDA and SIMCA were employed as supervised pattern recognition techniques for screening analysis.

LDA is focused on finding optimal boundaries between classes, maximizing between-class separability while minimizing within-class variability. However, LDA requires a selection of variables when full-spectrum data is available [45]. In SPA-LDA a training set consisting of n samples with known class labels is used to perform the process of variable selection. In the case of Raman spectroscopy, each sample consists of a spectrum with k wavenumbers. At first, the n training samples are centered on the mean of each class and stacked in the form of a **X** matrix ($n \times k$), where each column of **X** corresponds to a variable. Then, projection operations related to the columns of **X** are carried out to create k chains with l variables. Due to the loss of freedom degrees in the process of calculating class means, the chain length is limited by $n - c$, where c is the number of classes involved in the problem. Each time, the chain is initialized by one of the available k variables. Subsequent variables are selected to the chain in order to display the least collinearity with the previous ones, which is evaluated by the correlation between the respective column vectors of **X**. Finally, a total of $k \times l$ subsets of variables can be generated. These candidate subsets are assessed in terms of a cost function involving the average risk of misclassification over the validation set [33]. On this form, the optimal number of variables (NVs) is established for SPA-LDA model.

SIMCA is based on the advantage of soft modeling where each class can be independently modeled. This classification technique is based on producing a PCA model for each class of sample and then comparing their distance to the class confidence limits. On this form, SIMCA allows the analysis of a large data set concentrating the variance in few variables called principal components (PCs). Considering a set of n objects (samples) measured in k sensors (wavenumbers in Raman spectroscopy) generating the **X** matrix ($n \times k$), PCA reduced the matrix **X** into a product of two other arrays of low dimensionality **T** (n x A) and **L** (k x A) called scores and loadings, respectively, where A represents the number of new variables considered to be significant for the model of each class. The new variables in **T** present the advantage of being mutually orthogonal allowing the use of all spectral information in the construction of the SIMCA model. This characteristic permits the detection of anomalous samples or outliers, present in the data set. Finally, the classification of new samples is carried out by means of an F-test at a given significance level. The optimal number of PC's can be selected by leave-one-out cross-validation using the training set by means of the calculation of the root mean square error of cross validation (RMSECV) [30].

The optimal NVs and PCs were established for SPA-LDA and SIMCA, respectively. The performance of the classification models was evaluated by accuracy, which is defined using the ratio of samples in the test set correctly assigned into their respective classes.

2.6.3 Multivariate calibration model. Regression analysis was performed by applying the PLS algorithm. PLS is a factor analysis method which the full-spectrum advantage is retained by forming a new coordinate system, as follow:

$$\mathbf{X} = \mathbf{T} \times \mathbf{L} + \mathbf{E} \qquad \text{Eq. 2}$$

where **X** is the m x n matrix of calibration spectra, **L** is a h x n matrix with the rows of **L** being the new PLS basis set of h full-spectrum vectors, often called loading vectors or loading spectra. **T** is an m x h matrix of intensities (scores) in the new coordinate system of the h PLS loading vectors for the m sample spectra. **E** is now the $m \times n$ matrix of spectral residuals not fit by the optimal PLS model. The spectral intensities in the new coordinate system can be related to concentrations using least-squares analysis:

$$\mathbf{c} = \mathbf{T} \times \mathbf{b} + \mathbf{e} \qquad \text{Eq. 3}$$

where **b** is the $h \times l$ vector of coefficients relating the scores to the concentrations, **T** is the matrix of scores from the PLS spectral decomposition (Eq. 2) and **e** an error vector. The concentration in an unknown sample is obtained by using the previously calculated **b** coefficients [30].

In this work, the PLS model and its performance was evaluated by leaving one out cross validation, in which each sample of the calibration set was left out once, and its concentration was estimated by a model built with the remaining samples. The models were evaluated by calculating analytical parameters as the determination coefficient (R^2), root mean square error of prediction (RMSEP) and the relative error of prediction (REP%). In addition, figures of merit as sensitivity (SEN) was calculated as SEN = $1/||b_k||$ and limit of detection (LOD) and limit of quantification (LOQ) were calculated as LOD (or LOQ) = $k\ \alpha\ ||b_k||$, where k= 3.3 for LOD, and k= 10 for LOQ, α is the standard deviation of the net analytical signal (NAS), and $||b_k||$ is the Euclidean norm of the vector of regression coefficients estimated from the multivariate calibration model for analyte k.

Multivariate data analysis including KS, AsLS, SG, SPA-LDA, SIMCA and PLS algorithms was performed using MatLab R2018 software (The MathWorks, USA).

2.7 HPLC analysis

The quantification of OMC in the o/w microemulsions was performed by using a HPLC Agilent 1260 Infinity system equipped with an Eclipse Plus (Agilent Technologies) C18 reverse phase column (3.5 μm spherical particles covered with octadecylsilane, 100 × 4.6 mm). Methanol and ultrapure water in an 85:15 (v/v) ratio was the mobile phase and was controlled by the activation of quaternary pumps at 25 °C. The flow rate was 0.5 mL min^{-1}. OMC was monitored at 311 nm using a UV detector. Moreover, isocratic mode was employed and the injected sample volume was 20 μL. An external calibration curve was constructed using standard solutions of OMC in the range of 10-150 mg L^{-1}. Finally, the peak area obtained for a retention time of 16.5 min was used for the OMC quantification.

3. Results and discussion

3.1 Characterization of the OMC loaded microemulsions

As shown in Table 1, the Z values for all formulations were less than 100 nm, varying between 9.99 and 15.14 nm, in agreement with expected values for microemulsions [12]. In addition, the size was increasing as the concentration of OMC increased. The corresponding histograms are showing in Fig. S1. Thus, from the size distribution it is possible to obtain the PdI, which is the ratio between the standard deviation and the mean droplet size. PdI values between 0.1 and 0.25 indicate a narrow size distribution whereas a PdI>0.5 indicates a very broad size distribution. All the formulations presented a PdI between 0.082 and 0.286 (Table 1), indicating that the microemulsions tended to approach monodispersion and become stable systems due to the uniformity in the drops.

On the other hand, the conductivity of the microemulsions were in range 0.04-0.07 mS cm^{-1}, demonstrating that all the systems were o/w type. Furthermore, as expected, the values of conductivity decreased as the OMC concentration was higher (Table 1). All the OMC-MEs presented a pH range between 3.5 and 4.0 which is suitable for topical use [11]. Finally, as expected, the SPF values were increasing as the OMC concentration in the formulation increased (Table 1).

3.2 Band assignments

The Raman spectra of OMC, DME, OA and the o/w microemulsion are shown in Fig. 1 along with the chemical structures of the microemulsion components. As can be seen, the spectrum of OMC (green line) shows five strong peaks around 1171, 1204, 1605, 1635 and 1706 cm^{-1}. The peak at 1171 cm^{-1} is assigned to the C–H bend, while the mode at 1204 cm^{-1} correspond to the C–O–C bend of the *p*-methoxy group. The modes at 1605 cm^{-1} are assigned to aromatic ring vibrations. In addition, the peaks at 1706 and 1635 cm^{-1} are assigned to the conjugated ester C=O stretch and the conjugated exocyclic C=C stretch, respectively [46]. In the case of the oleic acid (red line) the spectrum shows peaks at 1655, 1439 and 1302 cm^{-1}. The peak at 1655 cm^{-1} can be attributed to the stretching C=O bonds in the carboxylic group. The peak at 1439 cm^{-1} is assigned to CH_2 scissoring mode while the peak at 1302 cm^{-1} is attributable to CH_2 wagging and the carboxylic C–O stretching mode. In addition, the peaks at 1274, and 972 cm^{-1} can be assigned to stretching C–O bonds, and OH out of plane bending modes, respectively [47]. Concerning the DME spectrum (blue line), it shows the main peaks around 1606, 1562, 1413, 1347, and 1040 cm^{-1}. The peaks at 1606, 1562 and 1413 cm^{-1} are attributed to the stretching C=C bonds while the O-H in plane bending appears at 1347 cm^{-1}. Further, the band at 1040 cm^{-1} correspond to the C–O stretching vibration [48]. Finally, the ME spectrum (black line) was similar to the surfactant spectrum with bands in 1606, 1562, 1413, 1347 and 1040 cm^{-1} due to the fact that the ME contains DME as principal component (21.9%).

3.3 Multivariate analysis

Before the multivariate analysis, the spectra data were preprocessed as described in section 2.6.1. Fig. S2a and S2b show the Raman spectra corresponding to the training and the prediction sets before and after the data preprocessing step using AsLS and SG algorithms, respectively. Moreover, an exploratory analysis of the data was performed using PCA. As can be seen in Fig. S3, the graph of scores (PC1 vs PC2) shows that discrimination of the samples was not achieved and the use of supervised pattern recognition techniques was required.

3.3.1 Multivariate classification models

The statistic results corresponding to SPA-LDA and SIMCA methods obtained for the classification of o/w microemulsion samples containing different concentrations of OMC are presented in Table 2.

In the case of SPA-LDA, the minimum value of the cost function (0.378) was obtained using seven spectral variables. Fig. 2 shows a Raman spectrum of an OMC loaded microemulsion (5.0%) indicating the seven selected variables by SPA: 1171, 1340, 1370, 1400, 1560, 1580 and 1605 cm^{-1}. As can be seen, several variables agree with the regions of higher intensity for OMC (1171, 1580 and 1605 cm^{-1}). However, some variables outside these regions are important and have also been identified by SPA. For example, the variables 1340, 1370, 1400 and 1560 corresponding to the surfactant (DME) were also selected. Then, a LDA model was obtained and the graph of the Fisher discriminant functions (DF) are showed in Fig. 3. As can be seen, the discrimination between ME/ME-1%OMC and ME-1%OMC/ME-3%OMC samples in DF1 is observed, whereas DF2 discriminated ME-7%OMC/ME-10%OMC samples. Likewise, ME-3%OMC samples were discriminated from ME-5%OMC samples and ME-5%OMC samples were discriminated from ME-7%OMC samples in DF4. In terms of performance, the SPA-LDA model achieved an accuracy of 97.8% since only one sample corresponding to ME-7%OMC class was misclassified.

On the other hand, for SIMCA model the optimal number of principal components for defining each class was four with an explained variance higher than 97% (α=0.05). As can be seen in the Coomans graph (Fig. S4), the model allowed the correct classification of all analyzed samples achieving an accuracy of 100%. However, it is important to note that SIMCA used a total of 930 spectral variables in contrast to SPA-LDA which used only 7, reducing in a 99.2% the number of variables used to the construction of the classification model.

3.3.2 Multivariate calibration model

PLS regression was performed using six latent variables (PCs) which explained 97.5% of the total variance. The analytical parameters (R^2, RMSEP and REP%) and figures of merit (SEN, LOD and LOQ) of the PLS calibration model are shown in the Table 3. The R^2 (0.9699) was appropriate and both RMSEP (0.5434% (w/w)) and REP (9.88%) values were reasonably low indicating that the proposed method is accurate in the prediction of samples. This fact is important since that the samples contain high concentrations of other components such as oil (OA), surfactant (DME) and cosurfactant (ET) which could interfere significantly. With regard to the figures of merit, satisfactory values for SEN (52.19% (w/w)), LOD (0.0300% (w/w)) and LOQ (0.0990% (w/w)) were obtained considering the OMC concentration present in the analyzed formulations. However, the LOD and LOQ values were higher than those obtained by other authors using separation techniques as HPLC [49-53].

3.4 Mapping analysis

The ultra-high resolution (both spectrally and spatially) of the Raman microscope provides information about the identity and presence of specific components and reveals their location and distribution within a sample area [29]. The proposed PLS-CRS method allowed obtaining distribution maps (2D image) which provided a visual representation of the OMC distribution in the analyzed samples. A rainbow color-coded composite image was developed to show the localization of OMC. Fig. 4 shows the color spatial mapping corresponding to the predicted distribution of OMC in formulations with different concentration of OMC (3.0, 5.0 and 7.0%). The pixel concentration was predicted by multiplying each Raman spectrum by the corresponding PLS regression vector. As can be seen, over an area of 204 μm^2, the pixels that present orange-yellow color intensity correspond to higher concentrations of OMC, showing in light blue-blue regions with lower concentrations of OMC. As expected, the orange-yellow color pattern represented a larger area as the concentration of OMC increased. In relation to distribution, the OMC seems non-homogeneously distributed in the microemulsion samples. It is possible that oil drops have different drug loading capacity, which is reflected in regions with different concentrations of OMC. It is important to note that CRS was appropriate to predict the content of OMC and to depict its distribution in samples containing different concentrations of the UV-filter.

3.5 Validation of CRS-PLS method

Validation of trueness was carried out by comparison of the results obtained by the proposed method with the ones obtained by the reference method (HPLC). The statistical comparison was carried out using the joint interval test for the slope and the intercept [54]. The values obtained by the proposed method were regressed against the HPLC method. The estimated intercept (a) and slope (b) obtained, were compared with the ideal values of intercept = 0 and slope = 1. The elliptical joint confidence regions obtained (Fig. S5) indicate that there are no significant statistical differences between the results obtained by both methods, considering an overall significance level of $\alpha = 0.05$. From these results it can be concluded that, in comparison to the reference method, there is no bias in the results obtained for OMC.

On the other hand, Table 4 compares the proposed method with respect to different HPLC methods [49-53] developed to quantify OMC in different PPCPs. As can be seen, the proposed CRS-PLS method does not require a separation technique or the use of organic solvents. Moreover, the analysis time is short mainly due to the fact that a rapid instrumental response is obtained using CRS in comparison to the HPLC methods.

3.6 Green approach

In order to assess the greenness of an analytical methodology, it is necessary to have metrics to evaluate and easily quantify this parameter. The AGREE metric [55] is a recently proposed method that provides an analytical greenness estimation calculator, based on the twelve principles of GAC. The Fig. S6a and S6b show the colored pictogram corresponding to the proposed green CRS and HPLC methods, respectively. As can be seen, a numerical value of 0.85 was obtained for the proposed method, which is close to the ideal value of 1.0, demonstrating the greenery of the method. This fact is relating to a non-destructive analysis of the sample, requiring a minimal volume of it for the complete analysis (10 μL). In addition, there was no need of any pretreatment or derivatization steps, which significantly reduces the total time of analysis. The procedure did not include reagents and organic solvents and no waste was generated, which is an important benefit in terms of safety for both, the operator and the environment. Since the waste generation is one of the most important points in the evaluation of greenness of an analytical method we assigned a higher weight value to this parameter. This fact can be observed in the pictogram in which all the points represent the same area with the exception of number seven which correspond to waste. On the other hand, the points 3, 5 and 8 were not green, indicating a certain distance from the ideal value that is 1.0. As can be seen in Fig. S7, the point 3 is related to: *If possible, the measurements should be performed in situ*. In our case, this point was the farthest from the ideal greenness value, since the measurement device was in off-line position. Although these measurements were carried out off-line smaller at-line Raman spectrometers could be used instead once analytical methods are established. Point 5 states: *Automated and miniaturized method should be selected*. In the pictogram this point was yellow representing a smaller distance from the ideal value since the proposed method did not include automation. Finally, point 8 is related to: *Multi-analyte or multi-parameter methods are preferred versus methods using one analyte at a time*. This point was distant from the ideal value for two reasons: firstly, because only one analyte was determined and after because the throughput was 15 h^{-1}. This value was relatively less because included the chemometric analysis step. On the contrary, the pictogram corresponding to the HPLC method shows that the overall value of greenness provided by the metric is much lower (0.47). This fact is mainly due to the use of organic solvents required by the chromatographic technique, the absence of automation and/or miniaturization, and the impossibility of performing in situ analysis, among others. In particular, the use of methanol negatively impacts the analysis since this solvent is not obtained from renewable sources and it is a potentially dangerous solvent for the operator and the environment (toxic and highly flammable). Moreover, the HPLC method generated waste which should be avoided (Fig. S7). In conclusion, the value obtained for the proposed CRS method is much higher than the HPLC method and it corresponds to an excellent greenness.

4. Conclusions

This work demonstrated that the combination of confocal Raman spectroscopy and chemometrics is a potential tool to both qualitative and quantitative analysis of OMC loaded in oil in water microemulsions, alternatively to the traditionally chromatographic methods. In addition, confocal Raman spectroscopy allowed direct analysis of samples containing high concentration of surfactant, cosurfactant and oil phase, avoiding pretreatment steps, and obtaining fast, simple and reliable analytical methods in concordance whit the principles of green analytical chemistry. The combination of confocal Raman spectroscopy and SPA-LDA and SIMCA provided accurate methods for discriminate different percentages of OMC in microemulsion samples demonstrating the potential use of this methodology as a screening tool. On the other hand, a PLS model based on the analysis of Raman spectra exhibited optimal analytical parameters for the quantification of OMC at different concentration levels which represents an important fact since the concentration of OMC is directly related to the solar protection factor and therefore with the efficacy of the formulation. In addition, the proposed PLS-CRS method allowed obtaining distribution maps which provided a visual representation of the OMC distribution in the analyzed samples. Finally, the use of green chemistry metrics allowed to demonstrate the greenery of the proposed method obtaining a high value, which was close to the corresponding an ideal green methodology.

Acknowledgments

The authors gratefully acknowledge the financial support of Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET; PIP 11220120100625), Ministerio de Ciencia y Tecnología (PICT, 201-0659) and Universidad Nacional del Sur (PGI 24Q086 and PGI 24/B272).

References

[1] S. Dey, F. Bano, A. Malik, Pharmaceuticals and personal care products (PPCP) contamination- a global discharge inventory, in: Butterworth-Heinemann (Eds.), Phamarceuticals and personal care products: Waste Management and Treatment Technology, Oxford, 2019, pp. 1-26.

[2] S.E. Mancebo, J.Y. Hu, S.Q. Wang, Sunscreens: A review of health benefits, regulations and controversies, *Dermatol. Clin.* 32 (2014) 427-438. http://dx.doi.org/10.1016/j.det.2014.03.011

[3] R. Jansen, U. Osterwalder, S.Q. Wang, M. Burnett, H.W. Lim, Photoprotection: part II. Sunscreen: development, efficacy, and controversies, J. Am. Acad. Dermatol. 69 (2013) 867.e1-867.e14. https://doi.org/10.1016/j.jaad.2013.08.022

[4] J.C. Cardoso, B.M.L. Armondes, J.B.G. Júnior, T.A. de Araújo, J.L.R. Júnior, N.R. Poppi, V.S. Ferreira, Determination of 4-methylbenzilidene camphor in sunscreen by square wave voltammetry in media of cationic surfactant, Microchem. J. 85 (2007) 301-307. https://doi.org/10.1016/j.microc.2006.07.006

[5] D.S. Nascimento, V. Volpe, M. Insausti, M. Grünhut, On line kinetic analysis of permeation profiles for UV filter loaded microemulsions using an automatic system with spectroscopic detection and a chemometric approach, RSC Adv. 11 (2021) 15528-15538. http://doi.org/10.1039/D1RA01067K

[6] G.J. Nohynek, H. Schaefer, Benefit and risk of organic ultraviolet filters, *Regul. Toxicol. Phamarcol.* 33 (2001) 285-299. http://doi.org/10.1006/rtph.2001.1476

[7] FDA, Department of Health and Human Services, 21 CFR Parts 310, 352, 700 and 740, RIN 0910-AA01, Sunscreen Drug Products for the Over-the-Counter Human Use Final Monograph, Federal Register, Rules and Regulations, 64 (1999) 27666-27693. https://www.govinfo.gov/content/pkg/FR-1999-05-21/pdf/99-12853.pdf

[8] EU, European Directive 76/768/EEC and its successive amendments, basic act 31976L0768. https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=celex:31976L0768

[9] D.S. Nascimento, M. Insausti, B.S.F. Band, M. Grünhut, Photolysis study of octyl p-methoxycinnamate loaded microemulsion by molecular fluorescence and chemometric approach, Spectrochim. Acta A. 191 (2018) 277-282. http://dx.doi.org/10.1016/j.saa.2017.10.025

[10] V. Volpe, D.S. Nascimento, M. Insausti, M. Grünhut, Octyl p-methoxycinnamate loaded microemulsion based on Ocimum basilicum essential oil. Characterization and analytical studies for potential cosmetic applications, Colloids Surf. A Physicochem. Eng. Asp. 546 (2018) 285-292. https://doi.org/10.1016/j.colsurfa.2018.02.070

[11] M. Fanun, Microemulsions as delivery systems, Curr. Opin. Colloid Interface Sci. 17 (2012) 306-313. https://doi.org/10.1016/j.cocis.2012.06.001

[12] V. Volpe, M.N. Giacomodonato, D.O. Sordelli, M. Insausti, F.R. Buzzola, M. Grünhut, Ciprofloxacin loaded o/w microemulsion against Staphylococcus aureus. Analytical and biological studies for topical and intranasal administration, J. Drug Deliv. Sci. Technol. 57 (2020) 101705. https://doi.org/10.1016/j.jddst.2020.101705

[13] S. Scalia, Determination of sunscreen agents in cosmetic products by supercritical fluid extraction and high-performance liquid chromatography, J. Chromatogr. A. 870 (2000) 199-205. http://doi.org/10.1016/s0021-9673(99)01224-8

[14] L. Rubio, L. Valverde-Som, L.A. Sarabia, M.C. Ortiz, Improvement in the identification and quantification of UV filters and additives in sunscreen cosmetic creams by gas chromatography/mass spectrometry through three-way calibration

techniques, Talanta. 205 (2019) 120156. https://doi.org/10.1016/j.talanta.2019.120156

[15] V.S. Ferreira, J.B.G. Júnior, C.M.S.C. Oliveira, R.M. Takeuchi, A.L. Santos, M.A.G. Trindade, Voltammetric analysis of sun-block preparations containing octocrylene and its association with 2-Hydroxy-4-methoxybenzophenone and octyl-methoxycinnamate, Microchem. J. 106 (2013) 378-383. https://doi.org/10.1016/j.microc.2012.10.002.

[16] A. Chisvert, M.T. Vidal, A. Salvador, Sequential injection analysis for benzophenone-4 and phenylbenzimidazole sulphonic acid in sunscreen sprays by solid-phase extraction coupled with ultraviolet spectrometry, Anal. Chim. Acta. 464 (2002) 295-301. https://doi.org/10.1016/S0003-2670(02)00432-4

[17] M. Haunschmidt, W. Buchberger, C.W. Klampfl, R. Hertsens, Identification and semi-quantitative analysis of parabens and UV filters in cosmetic products by direct-analysis-in-real-time mass spectrometry and gas chromatography with mass spectrometric detection, Anal. Methods. 3 (2011) 99-104. https://doi.org/10.1039/C0AY00588F

[18] K. Mori, K. Itoh, S. Suzuki, H. Nakamura, Analysis of ultraviolet absorbers in cosmetics by two dimension NMR spectroscopy, J. Toxicol. Environ. Health. 42 (1996) 60-66. http://doi.org/10.1248/jhs1956.42.60

[19] X.H. Liu, X.Z. Liang, X. Fang, W.P. Zhang, Preparation and evaluation of novel octylmethoxycinnamate- loaded solid lipid nanoparticles, Int. J. Cosmet. Sci. 37 (2015) 446-453. https://doi.org/10.1111/ics.12216

[20] S.A. Oladepo, G.R. Loppnow, Ultraviolet resonance Raman spectroscopy as a robust spectroscopic tool for in situ sunscreen analysis Anal. Chim. Acta. 628 (2008) 57-66. http://doi.org/10.1016/j.aca.2008.08.031

[21] M. Mei, X. Huang, Online analysis of five organic ultraviolet filters in environmental water samples using magnetism-enhanced monolith-based in-tube

solid phase microextraction coupled with high-performance liquid chromatography, J. Chromatogr. A. 1525 (2017) 1-9. https://doi.org/10.1016/j.chroma.2017.09.065

[22] M. Vila, J.P. Lamas, C. Garcia-Jares, T. Dagnac, M. Llompart, Ultrasound-assisted emulsification microextraction followed by gas chromatography-mass spectrometry and gas chromatography-tandem mass spectrometry for the analysis of UV filters in water, Microchem. J. 124 (2016) 530-539. https://doi.org/10.1016/j.microc.2015.09.023

[23] F.D. Fleischli, S. Mathes, C. Adlhart, Label free non-invasive imaging of topically applied actives in reconstructed human epidermis by confocal Raman spectroscopy, Vib. Spectrosc. 68 (2013) 29-33. http://doi.org/10.1016/j.vibspec.2013.05.003

[24] M. Sohn, T. Buehler, G. Imanidi, Repartition of oil miscible and water soluble UV filters in an applied sunscreen film determined by confocal Raman microspectroscopy, Photochem Photobiol Sci. 15 (2016) 861-871. https://doi.org/10.1039/C6PP00024J

[25] M.C. Breitkreitz, G.P. Sabin, G. Polla, R.J. Poppi, Characterization of semi-solid self-emulsifying drug delivery systems (SEDDS) of atorvastatin calcium by Raman image spectroscopy and chemometrics, J. Pharm. Biomed. Anal. 73 (2013) 3-12. http://doi.org/10.1016/j.jpba.2012.03.054

[26] E.M. Paiva, V.H. Silva, R.J. Poppi, C.F. Pereira, J.J.H. Rohwedder, Comparison of macro and micro Raman measurement for reliable quantitative analysis of pharmaceutical polymorphs J. Pharm. Biomed. Anal. 157 (2018) 107-115. http://doi.org/10.1016/j.jpba.2018.05.010

[27] D.J. Silva, H. Wiebeck, Using PLS, iPLS and siPLS linear regressions to determine the composition of LDPE/HDPE blends: A comparation between confocal Raman and ATR-FTIR spectrocopies, Vib. Spectrosc. 92 (2017) 259-266. https://doi.org/10.1016/j.vibspec.2017.08.009

[28] L.M. Duarte, D. Paschoal, C.M.S. Izumi, M.D. Dolzan, V.R. Alves, G.A. Micke, H.F. Santos, M.A.L. Oliveira, Simultaneous determination of aspartame, cyclamate,

saccharin and acesulfame-K in powder tabletop sweeteners by FT-Raman spectroscopy associated with the multivariate calibration: PLS, iPLS and siPLS models were compared, Int. Food Res. J. 99 (2017) 106-114. https://doi.org/10.1016/j.foodres.2017.05.006

[29] A.A. Saleh, M. Hegazy, S. Abbas, A. Elkosasy, Development of distribution maps of spectrally similar degradation products by Raman chemical imaging microscope coupled with a new variable selection technique and SIMCA classifier, Spectrochim. Acta A Mol. Biomol. Spectrosc. 268 (2022) 120654. https://doi.org/10.1016/j.saa.2021.120654

[30] K.R. Beebe, R.J. Pell, M.B. Seasholtz, Chemometrics: A Practical Guide. John Wiley & Sons, inc. 1998.

[31] C.M. Andersen, R. Bro, Variable selection in regression-A Tutorial, J. Chemom. 24 (2010) 728-737. https://doi.org/10.1002/cem.1360

[32] M.C.U. Araújo, T.C.B. Saldanha, R.K.H Galvão, T. Yoneyama, H.C. Chame, V. Visani, The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. Chemom. Intell. Lab. Syst. 57 (2001) 57-65. https://doi.org/10.1016/S0169

[33] M.J.C. Pontes, R.K.H. Galvão, M.C.U. Araujo, P.N.T. Moreira, O.D. Pessoa Neto, G.E. José, T.C.B. Saldanha, The successive projections algorithm for spectral variable selection in classification problems, Chemom. Intell. Lab. Syst. 78 (2005) 11-18. https://doi.org/10.1016/j.chemolab.2004.12.001

[34] C.C. Acebal, M. Grünhut, A.G. Lista, B.S.F. Band, Successive projections algorithm applied to spectral data for the simultaneous determination of flavour enhancers, Talanta 82 (2010) 222-226. https://doi.org/10.1016/j.talanta.2010.04.024

[35] K. Wang, D.W. Sun, Q. Wei, H. Pu, Quantification and visualization of α-tocopherol in oil-in-water emulsion based delivery systems by Raman microspectroscopy, LTW-Food Sci. Technol. 96 (2018) 66-74. https://doi.org/10.1016/j.lwt.2018.05.017

[36] P.T. Anastas, M.M. Kirchhoff, Origins, current status, and future challenges of green chemistry. Acc. Chem. Res. 35 (2002) 686-694. https://doi.org/10.1021/ar010065m

[37] M.F. Razuc, M. Grünhut, E. Saidman, M. Garrido, B.S.F. Band, Green method based on a flow-batch analyzer system for the simultaneous determination of ciprofloxacin and dexamethasone in pharmaceuticals using a chemometric approach, Talanta. 115 (2013) 314-322. https://doi.org/10.1016/j.talanta.2013.05.002

[38] M. Espino, F.J.V. Gómez, J. Boiteux, M.Á. Fernández, M.F. Silva, Green Chemistry Metrics, Comprehensive Foodomics. 2021, pp. 825-833. https://doi.org/10.1016/B978-0-08-100596-5.22822-9

[39] A. Gałuszka, Z.M. Migaszewski, P. Konieczka, J. Namiésnik, Analytical Eco-Scale for assessing the greenness of analytical procedures, TrAC Trends Analyt. Chem. 37, (2012) 61-72. https://doi.org/10.1016/j.trac.2012.03.013

[40] J.S. Mansur, M.N. Breder, M.C. Mansur, R.D. Azulay, Determination of sun protection factor by spectrophotometry, An. Bras. Dermatol. 61 (1986) 121-124.

[41] R.M. Sayre, P.P. Agin, G.J. Levee, E. Marlowe, A comparison of in vivo and in vitro testing of sunscreening formulas, Photochem. Photobiol. 29 (1979) 559-566. http://dx.doi.org/10.1111/j.1751-1097.1979.tb07090.x

[42] R.W. Kennard, L.A. Stone, Computer-aided design of experiments. Technometrics, 11 (1969) 137-148. https://doi.org/10.2307/1266770

[43] H. Shixuan, W. Zhang, L. Liu, L. Huang, J. He, W. Xie, P. Wu, C. Du, Baseline Correction for Raman Spectra Using Improved Asymmetric Least Squares, Anal. Methods. 6 (2014) 4402-4407. https://doi.org/10.1039/C4AY00068D

[44] A. Savitzky, M.J.E. Golay, Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Anal. Chem. 36 (1964) 1627-1639. https://pubs.acs.org/doi/pdf/10.1021/ac60214a047

[45] D.D. de Sousa Fernandes, V.E. de Almeida, M.M. Fontes, M.C.U. de Araújo, G. Véras, P.H.G.D. Diniz, Simultaneous identification of the wood types in aged cachaças

and their adulterations with wood extracts using digital images and SPA-LDA, Food Chem. 273 (2018) 77-84. https://doi.org/10.1016/j.foodchem.2018.02.035

[46] L. Beyere, S. Yarasi and G.R. Loppnow, Solvent effects on sunscreen active ingredients using Raman spectroscopy, J. Raman Spectrosc. 34 (2003) 743-750. https://doi.org/10.1002/jrs.1042

[47] W.N. Silverio, G.V.S. Dutra, A.T. Jensen, O.A. Araújo, V. Garg, A.C. Oliveira, L.F. Valadares, F.G. Souza, F. Machado, Superparamagnetic nanoparticles stabilized with free-radical polymerizable oleic acid-based coating, J. Alloy Compd. 739 (2018) 1025-1036. https://doi.org/10.1016/j.jallcom.2017.12.338

[48] S. Sun, R.L. Birke, J.R. Lombardi, Surface-Enhanced Raman Spectroscopy of Surfactants on Silver Electrodes, J. Phys. Chem. 94 (1990) 2005-2010. https://doi.org/10.1021/j100368a052

[49] L. Montenegro, C. Carbone, G. Puglisi, Vehicle effects on in vitro release and skin permeation of octylmethoxycinnamate from microemulsions, Int. J. Pharm. 405 (2011) 162-168. https://doi.org/10.1016/j.ijpharm.2010.11.036

[50] E. Ramón, C. Alonso, L. Coderch, A. de la Maza, O. López, J.L. Parra, J. Notario, Liposomes as Alternative Vehicles for Sun Filter Formulations, Drug Deliv. 12 (2005) 83-88. https://doi.org/10.1080/10717540490446080

[51] M.S. Monteiro, R.A. Ozzetti, A.L. Vergnanini, L. de Brito-Gitirana, N.M. Volpato, Z.M. de Freitas, E. Ricci-Júnior, E.P. dos Santos, Evaluation of octyl p-methoxycinnamate included in liposomes and cyclodextrins in anti-solar preparations: preparations, characterizations and in vitro penetration studies, Int. J. Nanomedicine. 7 (2012) 3045-3058. http://doi.org/10.2147/IJN.S28550

[52] M. Wharton, M. Geary, N. O'Connor, B. Murphy, A rapid high performance liquid chromatographic (HPLC) method for the simultaneous determination of seven UV filters found in sunscreen and cosmetics, Int. J. Cosmet. Sci. 33 (2011) 164-170. https://doi.org/10.1111/j.1468-2494.2010.00607.x

[53] J. Teixeira, Z.M.F. de Freitas, M.S. de Monteiro, A.B. Vermelho, E. Ricci Junior, E.P. dos Santos, Development and characterization of photoprotective formulations containing keratin particles, Braz. J. Pharm. Sci. 55 (2019) 17649. https://doi.org/10.1590/s2175-97902019000217649

[54] N. Draper, H. Smith, Applied Regression Analysis, 2° Ed., John Wiley and Sons, New York, 1981.

[55] F. Pena-Pereira, W. Wojnowski, M.Tobiszewski, AGREE-Analytical GREEnness Metric Approach and Software, Anal. Chem. 92 (2020) 10076-10082. https://dx.doi.org/10.1021/acs.analchem.0c01887

Tables

Table 1. Physicochemical characterization of the analyzed OMC loaded o/w microemulsions (1.0; 3.0; 5.0; 7.0 and 10%). Z: average droplet size; PdI: polydispersity index; C: conductivity; SPF: solar protection factor. All measurements were performed in triplicate.

Sample	Z (nm)	PdI	C (mS cm^{-1})	pH	SPF
ME-1%OMC	9.99 ± 0.04	0.101 ± 0.070	0.070 ± 0.002	4.10 ± 0.01	1.73 ± 0.04
ME-3%OMC	9.93 ± 0.11	0.086 ± 0.045	0.060 ± 0.001	4.10 ± 0.01	3.92 ± 0.03
ME-5%OMC	10.25 ± 0.05	0.286 ± 0.046	0.050 ± 0.001	4.20 ± 0.03	4.12 ± 0.01
ME-7%OMC	13.71 ± 0.23	0.281 ± 0.071	0.050 ± 0.002	4.00 ± 0.01	5.13 ± 0.04
ME-10%OMC	15.14 ± 1.36	0.285 ± 0.018	0.040 ± 0.001	4.10 ± 0.02	18.20 ± 0.07

Table 2. Summary of the statistical results (95% confidence level) for SPA-LDA and SIMCA models applied to confocal Raman spectra data for the classification of o/w microemulsion samples containing different concentration of OMC.

Sample data set	SPA-LDA			SIMCA		
	NVs	Error	Accuracy (%)	PCs	Error	Accuracy (%)
Training (N=30)	7	0	100	4	0	100
Prediction (N=15)		1	97.8		0	100

NVs: number of variables; PCs: principal components

Table 3. Analytical parameters of the PLS calibration model for OMC determination in microemulsion samples.

Analytical parameter	Value
Spectral range (cm^{-1})	600-1700
Concentration range (% w/w)	0-10
Latent variables	6
R^2	0.9699
RMSEP (% w/w)	0.5434
REP (%)	9.88
SEN (% w/w)	52.19
LOD (% w/w)	0.0300
LOQ (% w/w)	0.0990

Table 4. Comparison of methods developed to determination of OMC in different PPCPs.

Method	Separation technique	Detection technique	Sample	Solvent	Time analysis (s)
[45]	HPLC	UVS (310 nm)	o/w microemulsion	ACN : Water (80:20; v/v)	ND
[46]	HPLC	UVS (307 nm)	o/w emulsion	Water : MeOH (12:88; v/v)	330
[47]	HPLC	UVS (310 nm)	Gel cream	Water : MeOH : (13:87; v/v)	ND
[48]	HPLC	UVS (313 nm)	Sunscreen/Cosmetics	EtOH : Acetic Acid 1% (70:30; v/v)	420
[49]	HPLC	UVS (320 nm)	Sunscreen	ACN:Water (90:10; v/v)	600
This work	Not used	CRS (600-1700 cm^{-1})	o/w microemulsion	Not used	20

UVS: UV spectroscopy; ND: not declared

Figure captions

Figure 1. Raman spectra corresponding to OMC-ME (black line), DME (blue line), OA (red line) and OMC (green line). Wavenumbers (cm^{-1}) of the principal peaks are indicated along with the chemical structure of the components. The spectra have been shifted along the y-axis for a better comprehension.

Figure 2. Raman spectrum corresponding to a OMC loaded microemulsion (5.0%). The circles in black indicates the seven variables selected by SPA algorithm (1171, 1340, 1370, 1400, 1560, 1580 and 1605 cm^{-1}).

Figure 3. Fisher's discriminant functions (DF) obtained by SPA-LDA applied to Raman data set. The samples are represented as ME (blue circle), ME-1%OMC (red asterisk), ME-3%OMC (black diamond), ME-5%OMC (yellow square), ME-7%OMC (green star) and ME-10%OMC (violet cross).

Figure 4. Distribution maps representing the OMC predicted concentration by PLS regression in different o/w microemulsion samples (ME-1.0%OMC; ME-3.0%OMC; ME-7.0%OMC). The pixels that present highest colour intensity (orange-yellow) correspond to higher concentrations of OMC.

Figure 1

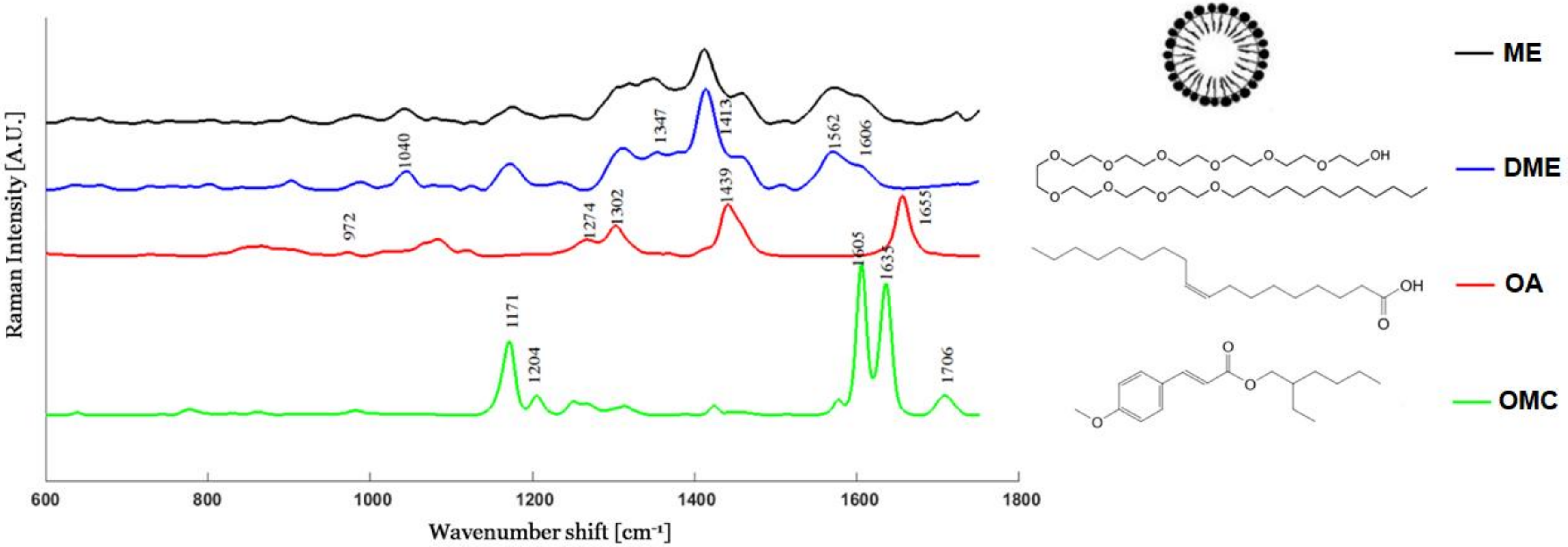

Raman Intensity [A.U.]
Wavenumber shift [cm⁻¹]
600
800
1000
1200
1400
1600
1800
1040
1347
1413
1562
1606
972
1274
1302
1439
1655
1171
1204
1605
1635
1706
OH
OH
O
O
O
ME
DME
OA
OMC

Figure 2

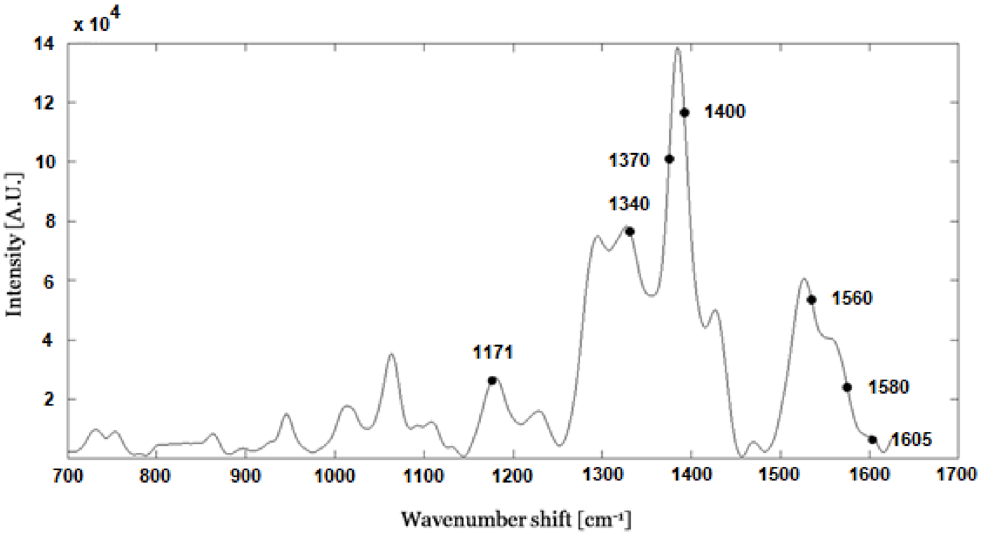
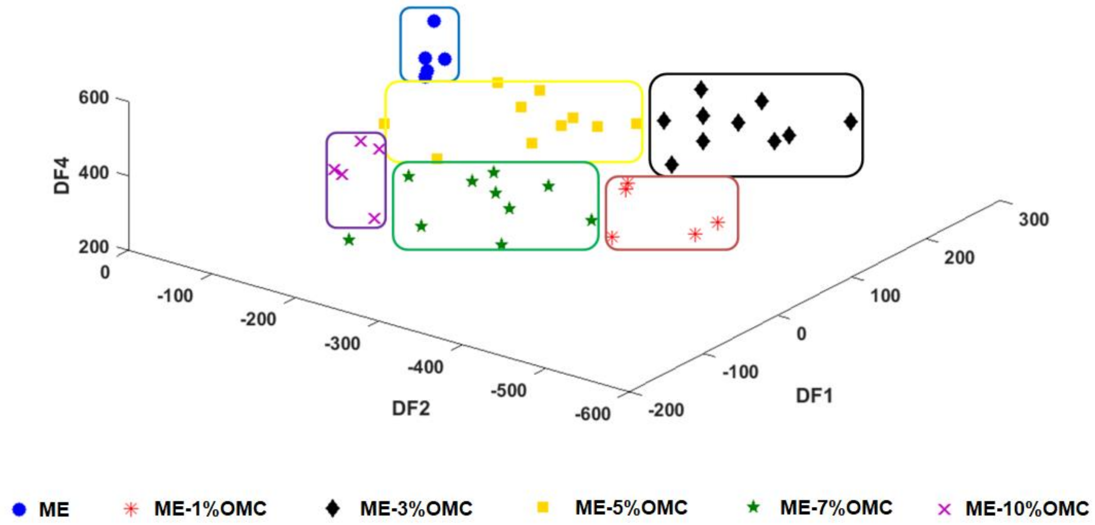
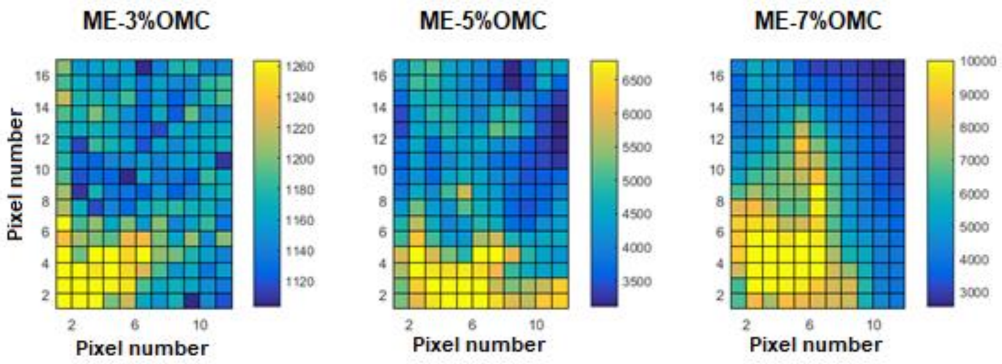

x 10^4
Intensity [A.U.]
14
12
10
8
6
4
2
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
Wavenumber shift [cm⁻¹]
1171
1340
1370
1400
1560
1580
1605

Figure 3



DF4
600
400
200
0
-100
-200
-300
-400
-500
-600
DF2
-200
-100
0
100
200
300
DF1
● ME
✳ ME-1%OMC
◆ ME-3%OMC
■ ME-5%OMC
★ ME-7%OMC
× ME-10%OMC

Figure 4


ME-3%OMC
ME-5%OMC
ME-7%OMC
Pixel number
Pixel number
16
14
12
10
8
6
4
2
2
6
10
1260
1240
1220
1200
1180
1160
1140
1120
6500
6000
5500
5000
4500
4000
3500
10000
9000
8000
7000
6000
5000
4000
3000